# ANÁLISIS CUALITATIVO COMPARATIVO TOMO II

## DIFUSO (fsQCA)

## Teoría y Práctica

Juan Mejía Trejo



# ANÁLISIS CUALITATIVO COMPARATIVO

# TOMO II

# DIFUSO (fsQCA)

## Teoría y Práctica

Juan Mejía Trejo



Este libro fue sometido a un proceso de dictamen por pares de acuerdo con las normas establecidas por el Comité Editorial de la Academia Mexicana de Investigación y Docencia en Innovación (AMIDI)

Primera edición, 2023


Av. Paseo de los Virreyes 920.
Col. Virreyes Residencial
C.P. 45110, Zapopan, Jalisco, México

**ISBN Tomo II: 978-607-59567-5-6**
**ISBN Obra Completa: 978-607-59567-0-1**

# Índice

**CAPÍTULO 6.**

# Introducción

El presente documento, es continuación de la obra, *Análisis Cualitativo Comparativo Nítido (***csQCA***) y su relación con la Innovación. Discusión e interpretación de resultados* y tiene por objetivo introducir al investigador en los conceptos y aplicaciones que conlleva el caso del análisis cualitativo comparativo (**QCA.***Qualitative Comparative Analysis*) con el conjunto de *datos difusos* (*fuzzy-sets.***fsQCA**) a través de demostraciones de su uso con ejemplos varios en la administración de la innovación.

Desde fines de los años 80, su precursor, el análisis cualitativo comparativo (**QCA**.*Qualitative Comparative Analysis*) ha estado en el centro de atención de la metodología de las ciencias sociales ya que se fundamenta, en el conjunto de relaciones y objetivos establecidos del descubrimiento de condiciones *suficientes* y *necesarias*.

Cabe señalar, que es común referirse a la versión Booleana original del **QCA** como **csQCA**, donde **cs** (*crisp-sets*) denota un conjunto de *datos nítidos*, la versión que permite utilizar las condiciones de múltiples categorías, se denomina como **mvQCA**, donde **mv** (*multi-value*) describe el valor múltiple y para la versión **fsQCA** donde **fs** (*fuzzy-set*) describe al conjunto de *datos difuso*. El objetivo del análisis de **QCA**, en general, es dar cuenta de un resultado particular, contra los análisis basados en *regresión*, que por el contrario, generalmente tienen como objetivo ser herramientas base para *explicar los efectos de las causas* **(Wagemann y Schneider, 2010)**.

Sobre el estado de las ciencias sociales, como base general de la administración de la innovación, **Sartori (1970)**, llegó a afirmar:

> *"...se aprecia un pésimo estado de la ciencia… oscilando entre dos extremos poco sólidos: pensamiento inconsciente, lo que hace una abrumadora mayoría, y un pensamiento demasiado consciente, hecho por una pequeña minoría…"*

Llamando a los estudiosos a adquirir capacitación en lógica (*primaria*):

> *"...para dirigir un curso intermedio entre mal manejo lógico crudo por un lado, y perfeccionismo lógico (y parálisis) por otro lado..."*

Fue a fines de la década de 1980, cuando el Dr. Charles Ragin trajo el álgebra Booleana y la teoría de conjuntos, para las ciencias sociales con su innovador libro *The Comparative Method* **(Ragin, 1987)** que describe a profundidad todo lo relativo a los componentes del **QCA**. Aún así, el verdadero estímulo en la atención comenzó algunos años después, con el libro *Fuzzy-sets Social Science* **(Ragin, 2000)**. Por ahora, ya hay académicos que usan los *métodos comparativos configuracionales,* dada la posibilidad en formalizar el *análisis orientado a casos* y de este modo, ofrecer herramientas para mejorar la investigación comparativa. Estos métodos son particularmente aptos para identificar lo *mínimamente necesario* y/o *mínimamente suficiente* (*combinaciones de*) condiciones que producen un *resultado de interés* (es decir, *evaluar las causas de los efectos*), con gran potencial de aplicaciones cuantitativas como en la ingeniería **(Mendel y Korjani, 2010; Marks, et al., 2018)**.

Dadas las ventajas que se ofrecen, en discusiones recientes sobre los métodos comparativos *configuracionales*, los estudiosos sostienen que los enfoques cuantitativos basados en regresiones vs. el **QCA**, se aplican mejor junto a otro **(Ragin, 2008; Schneider y Wagemann 2010; Rihoux, 2006)**. Sin embargo, existe una advertencia para los entusiastas, de que los académicos no deban convertirse en *monomaníacos* del **QCA (Ragin y Rihoux 2004, p. 6)**.

Por otro lado, se tienen primeros esfuerzos de la aplicación de **fsQCA,** en el área del emprendimiento y la innovación como el trabajo de **Kraus (et al., 2017)** donde se hace una recopliación de **77** artículos publicados de 2005 a 2016 con las palabras clave de: **fsQCA**, administración de negocios, empredimiento e innovación, revelando un incremento paulatino en estos campos para el uso de la **fsQCA**. Es así, que las ciencias de la administración orientadas a la innovación, tienen la posibilidad de aprovechar lo

realizado en las ciencias sociales a través de las importantes aportaciones del Dr. Ragin, por lo que la presente obra, se compone de once capítulos, los cuales describimos brevemente:

**CAPÍTULO 4. Análisis Cualitativo Comparativo de Datos Difusos (fsQCA).** Este capítulo reporta al lector un tipo de datos muy especial, el de tipo difuso que usa el software **fsQCA**, revelando conceptos básicos de su uso a través del conocimiento de su naturaleza, la posibilidad de usarlos por niveles y de manera continua. Se presentan operaciones del conjunto de *datos difusos*, tales como: la negación Booleana, la conjunción (lógica **AND**), la unión (lógica **OR**). Incluye modelos y métodos de calibración análisis de condiciones necesarias y suficientes así como de consistencia, cómo realizar tablas de verdad y analizar las esquinas de espacio vectorial. Se aplican catorce ejercicios que demuestran el uso del software **fsQCA** que confirman el uso de los principales conceptos.

**CAPÍTULO 5. Evaluación del conjunto de datos fsQCA.** El capítulo hace una descrpición de la importancia de lo que se conoce como condiciones **INUS** y **SUIN** así como de *consistencia y cobertura* en general. Se presenta el desarrollo manual de cinco casos que confirman el uso de los principales conceptos.

**CAPÍTULO 6. La calibración y su importancia en fsQCA**. Este capítulo abre un debate sobre las implicaciones de la calibración antes de iniciar mediciones en alcance y contexto, los pros y contras de utilizar indicadores por parte de la investigación cuantitativa, relacion de **SEM** vs. **fsQCA**, la investigación cualitativa y la necesidad de la calibración, los métodos directo e indirecto de calibración. Se presenta el desarrollo manual de cuatro casos que confirman el uso de los principales conceptos.

**CAPÍTULO 7. Pensamiento configuracional en fsQCA.** Este capítulo demuestra, la importancia del pensamiento *configuracional* desde el punto de vista cualitativo y cuantitativo, se discute sobre la evaluación del grado de pertenencia en una configuración así como la comparación de trayectorias causales.

**CAPÍTULO 8. Efectos netos en fsQCA**. El capítulo describe lo que se debe entender por *efectos netos*, los problemas que se tiene del mismo,

se propone un cambio de enfoque a tipos de casos, se hace exposición de comparación de casos de *datos difusos* y el análisis de *configuraciones*.

**CAPÍTULO 9. Efectos netos vs. Configuraciones en fsQCA.** A través de dos casos hipotéticos, se realiza el estudio y se desarrollan cada uno de los conceptos vistos explicando por comparación sus resultados, a fin de comprender la naturaleza de los efectos netos de los de configuración por **fsQCA**.

**CAPÍTULO 10. Guía de análisis csQCA.** Con datos *nítidos*, de un caso de innovación, se ofrece al lector una guía rápida y visual, a través de *screen-shots* y 8 ejercicios, cómo resolver un caso a través de realizar: análisis de condiciones necesarias, representar datos en la tabla de verdad, identificar contradictorios y remanentes lógicos, análisis de suficiencia, minimización de la tabla de verdad para ocurrencia y no ocurrencia del fenómeno e interpretación de resultados.

**CAPÍTULO 11. Guía de análisis fsQCA.** Con datos *difusos*, de un caso de innovación, se ofrece al lector una guía rápida y visual, a través de *screen-shots* y 9 ejercicios, cómo resolver un caso a través de realizar: acceso al archivo de trabajo, análisis de condiciones necesarias con ocurrencia del fenómeno y sus reportes, interpretación de la cobertura y la cobertura, generación de la tabla de verdad, la aplicación de la opción Standard Analyses, interpretación de las soluciones generadas, generación de la tabla de verdad con no ocurrencia del fenómeno, análisis de implicaciones principales, e interpretación del caso.

**GLOSARIO.** Se presenta un glosario de los términos más utilizados al respecto.

CAPÍTULO 4.

# Análisis Cualitativo Comparativo de Datos Difusos (fsQCA)

El manejo de estos temas siempre será una labor arriesgada ya que la comunidad principalmente académica, está acostumbrada a recurrir a conceptos cuantitativamente más controlados y *pareados*, por ejemplo, **Corcuf (2013, p. 19)**, anota:

> *"...ya desde sus comienzos, las ciencias sociales lidian con una serie de pares conceptuales, como material/ideal, objetivo/subjetivo, colectivo/individual..."*

Lo anterior, nos lleva a visualizar un mundo dicotómico que invita a elegir el uno por el otro **(Bendix y Berger, 1959)**. La dualidad aristotélica del tipo *todos/ninguno* establece *pertenencias nítidas a conjuntos dicotómicos*, cuya aplicación es habitual en lo que respecta a saber *qué es la innovación* y lo qué no lo es; qué es una *innovación del modelo de negocios* y qué no lo es; qué es una *crisis por falta de innovación* y qué no lo es; qué es *una política pública de innovación* y qué no lo es; y así un largo etcétera. Ese es el principio de dicotomización de la técnica **csQCA. Sartori (2011),** urge a los *comparativistas a seguir una serie de reglas para el análisis de conceptos*, recomendando a los investigadores que *los conceptos empíricos, es decir, sujetos al análisis y la comparación, no sean ambiguos ni vagos*. Los conceptos, siempre deberán alcanzar la claridad y corregir equívocos **(Sartori, 2011, pp. 199-210)** para ser usados *unívoca y coherentemente*. La apuesta por la claridad permite una mejor definición de conceptos en tanto que sus límites, componentes y puntos de corte quedan perfectamente determinados. **Sartori (2011)** admite que *no existe una solución general al problema de la línea de frontera*, en el caso de su ejemplo, señalar con

precisión *quién forma parte de la élite y quién no*, pero concluye que *el problema de la indefinición de los componentes se puede afrontar mucho antes que el del punto de corte* **(Sartori, 2011, p. 221).** Lo que propone es, por lo tanto, una pauta de construcción de conceptos basada en la *determinación absoluta de sus componentes claves*. Así también, construir *conceptos que puedan convertirse en instrumentos de trabajo a los que el investigador pueda recurrir eficientemente durante la recolección de datos*. En tal caso los conceptos *deben esquivar la abstracción* hacia una inclusión universal. Por lo que, en este contexto, se debe volver a la *definición de los elementos clave de los conceptos* antes que afrontar la incertidumbre de *decidir qué es cada cosa*.

No es siempre una tarea sencilla, dictaminar la verdad o falsedad tajante de un concepto o de una proposición. La localización de un *punto de corte* (*cross-overpoint*), *o frontera,* que sirva como separador entre elementos no se presenta siempre de forma inmediata e inequívoca. *Hay términos imprecisos y expresiones vagas que requieren de una teoría no estándar de conjuntos nítidos* (*crisp*). Expresiones frecuentes nos demuestran la *existencia de términos imprecisos* o de ciertos continuos entre dos elementos, como: una *empresa parcialmente equipada con tecnología innovadora, un proceso de innovación a medio hacer*, una *innovación de servicio altamente mejorable*, *una innovación a punto de concebirse*, etcétera. También es frecuente en ciencias sociales toparse con *conceptos adjetivados* de tal forma que nos referimos a *innovaciones imperfectas*, o *invenciones de bajo impacto, sistemas de innovación moderadas,* etcétera. Nótese que aquí estamos lidiando con *aproximaciones* más que con elementos *exactos*, por lo que refiriéndonos a **Zadeh (1992, p. 52)** indica que en la lógica difusa, *el razonamiento exacto se visualiza como un caso límite del razonamiento aproximado,* es decir, *que en lógica difusa todo es cuestión de grado (de pertenencia) debido a la existencia de predicados vagos*. Pero, ¿qué es, entonces, un *predicado (o condición) vago*? Según **Trillas (1992, p. 1)**, *es un predicado* **P** *(el nombre de la propiedad de los objetos) que, al aplicarlo a una cierta colección* **U,** *de objetos (el universo del discurso), esta no queda completamente clasificada en solo dos subclases*.

Para la lógica clásica, los límites de la vaguedad se establecen en términos de **(0)** y **(1)**, donde **(0)** representa la *no pertenencia* y **(1)** la *pertenencia completa*. La *lógica difusa*, propone que es preciso dar acomodo a casos que por su naturaleza o por la naturaleza de los conjuntos, quedan en una *posición intermedia*. Para algunas investigaciones puede convertirse en un *serio problema* plantear soluciones *dicotómicas*. Sin embargo, es importante *no caer en el error de confundir vaguedad con ambigüedad* y, con ello, malinterpretar el potencial y la aplicación de la *lógica difusa*. Así, para *una palabra o expresión es usada de forma ambigua si tiene más de un significado posible en un contexto determinado y si no es claro qué significado es el adecuado* **(Terricabras, 1992, p. 19)**. Ejemplo de ambigüedad lo encontramos en el adjetivo *muy innovador* ¿qué tan mucho o poco corresponde a ser innovador?

En estos casos, la dificultad para interpretar el significado de *mucho o muy no es resultado de su ambigüedad sino, de lo contrario, de la falta de un criterio absolutamente preciso y automático que establezca una cifra concreta que lo pueda medir*. *La vaguedad no es un defecto* ya que una *zona difusa* entre dos conjuntos bien definidos reporta riqueza cuando se trata de explotar un caso. Esta *singularidad intermedia,* es una mejor estrategia que intentar presionar de forma improcedente para que esta quepa dentro de un conjunto al que no pertenece. Para incorporar la *borrosidad* a nuestros análisis *debemos recurrir a lógicas y afirmaciones diferentes a las clásicas*, que asigna proposiciones solo a los valores de verdad **(1)** y **(0),** y la *lógica polivalente* que asigna valores de certeza variando entre **(0)** y **(1)** **(Trillas, 1992, p. 10)**. La *lógica difusa permite convertir los predicados en subconjuntos difusos del intervalo* **(0, 1)**. De tal forma, el supuesto de que un *predicado es verdadero* no responde a si es fácilmente reconocible como **(0)** o **(1)** sino que se deberán aplicar modificadores para verificar, o validar, si cumple con la condición de pertenecer, más o menos, a un *subconjunto*. Es decir, en relación con el ejemplo de *ingreso por innovación* que recibe una empresa, lo necesario para contextualizar si es *mucho* sería definir los dos subconjuntos de empresas que reciben más ingresos y los que reciben menos, para después *asignar grados de pertenencia a esos subconjuntos*.

De tal forma sabremos a qué subconjunto la empresa se encuentra más próximo y tendremos mejores instrumentos para la interpretación.

No obstante, a menudo el criterio que se emplea para definir fronteras entre subconjuntos extremos/polares es *contraproducente, subjetivo y arbitrario* si no contamos con *recursos teóricos y conceptuales para* ello. Tomando como referencia la *innovación incremental*, ¿en qué momento podemos considerar que una producto la tiene alta o baja? Si nuestra investigación versa sobre el *impacto social de la innovación*, ¿cuánto impacto social y de qué categoría es necesaria para considerarla una innovación con alcance masivo? Y para estudios sobre *innovación disruptiva*, ¿es posible establecer un momento preciso en el que una innovación de producto/servicio, debe ser reconocido como disruptivo? Para estas y otras muchas preguntas *no es fácil recurrir a respuestas nítidas y dicotómicas*. *Será tarea del investigador definir las fronteras difusas entre conjuntos con la ayuda de argumentos teóricos comparables*. No se trata de ofrecer una medición empírica precisa sino de *definir el alcance de la singularidad de los conceptos*. Por ello las investigaciones que siguen un enfoque de *conjuntos difusos* requieren, como sucede para el resto de las técnicas configuracionales, del constante diálogo entre teoría, tratamiento y recolección de los datos, e interpretación de los resultados.

Como se aprecia, la *lógica difusa* encuentra aplicaciones lingüísticas allí donde *la exactitud numérica y la lógica clásica enfrentan dificultades*. **Trillas et al. (1995)** ofrecen un argumento definitivo a favor de la lógica difusa. Los autores argumentan que *esta lógica intenta aportar unos modelos que permitan avanzar en el estudio tanto de los predicados vagos como de las formas de razonamiento usuales* **(Trillas et al., 1995, p. 128)**. Su aplicación es extensa y no solo se restringe a las ciencias sociales. De hecho, su promotor fue un matemático **(Zadeh, 1968)** que buscaba aplicaciones prácticas en la ingeniería.

Es preciso constatar que la *lógica difusa no busca reformular la construcción semántica de los conceptos como tampoco abre una vía hacia la malformación* (estiramiento conceptual). La *lógica difusa* ,advierte de que incluso conceptos perfectamente definidos tienen problemas para llevar de

un país a otro por el simple hecho de *carecer de referencias relativas que apoyen su aplicación práctica*. No se trata de construir definiciones conceptuales arriesgadamente inclusivas o intencionadamente inútiles. Tampoco se trata de construir conceptos absurdos que engloben a cosas tan diferentes con relaciones imposibles.

Muchos de los fenómenos que interesan a los científicos sociales con orientación a la administración sobre la actividad de la innovación, varían según el nivel de alcance y profundidad con los que se quiera aplicar. Por ejemplo, si bien está claro que existen CEOs de pymes que practican actividades de innovación también existen otras, que no lo hacen y por lo tanto, aquellas que se ubican como casos intermedios. Estos CEOs de pymes intermedias ni son completamente clasificadas con prácticas de actividad innovadora pero tampoco están excluidas por completo. El hecho de que muchos de los objetos de estudio de interés para los científicos sociales con orientación a la administración de la innovación no se ajusten de forma perfecta a los conjuntos de datos nítidos, puede parecer, que anulen todas las buenas razones esbozadas en el capítulo 1, para analizar el fenómeno social de la administración de la innovación, en términos de conjunto de datos establecidos (*set relations*) .

Por ejemplo, ¿tiene sentido pensar que las organizaciones innovadoras son un subconjunto de organizaciones que tienen recursos tecnológicos si ambos (organizaciones innovadoras y recursos tecnológicos), son medidos con escalas de intervalo muy finas y detalladas? Una razón, por la cual los científicos sociales con orientación a la administración de la innovación, llegan a ser reacios a estudiar los fenómenos de dicho objeto de estudio, en términos de conjunto de datos establecidos (*set relations*) es porque piensan que el estudio del conjunto de datos establecidos (*set relations*) está restringido a medidas de escala nominal. Estas escalas no solo se consideran primitivas, sino también las escalas de intervalo y de razón, que se hayan recodificado a escalas nominales (y, por lo tanto, degradado) son casi siempre sospechosas de escasez de valor (**Mejía-Trejo, 2019a,** capítulo 2).

Afortunadamente, un sistema matemático bien desarrollado está disponible para abordar el grado de pertenencia en conjuntos: la teoría de conjuntos difusos **(Zadeh, 1972, p. 1965).** Los conjuntos de *datos difusos*, son especialmente potentes porque permiten a los investigadores *calibrar* la pertenencia parcial del conjunto de datos difusas, utilizando valores en el intervalo entre **0.0** (ausencia, no pertenencia) y **1.0** (presencia, pertenencia completa) sin abandonar los principios establecidos en el marco teórico de la investigación y la operacionalización de las variables. Como se explicó en el capítulo 1, el conjunto de datos establecidos (*set relations*) son centrales al marco teórico de las ciencias sociales, sin embargo, la evaluación del conjunto de datos establecidos (*set relations*) está fuera del alcance de los métodos correlacionales convencionales **(Ragin, 2008)**.

## La naturaleza del conjunto de datos difusos (*fuzzy-sets*)

La propuesta de **Ragin (2000, 2008)** considera varios postulados relevantes de la lógica difusa y los adapta al procedimiento **QCA** para que puedan analizarse desde la metodología de la *teoría de conjuntos*. Esto se logra atendiendo a dos planteamientos singulares:

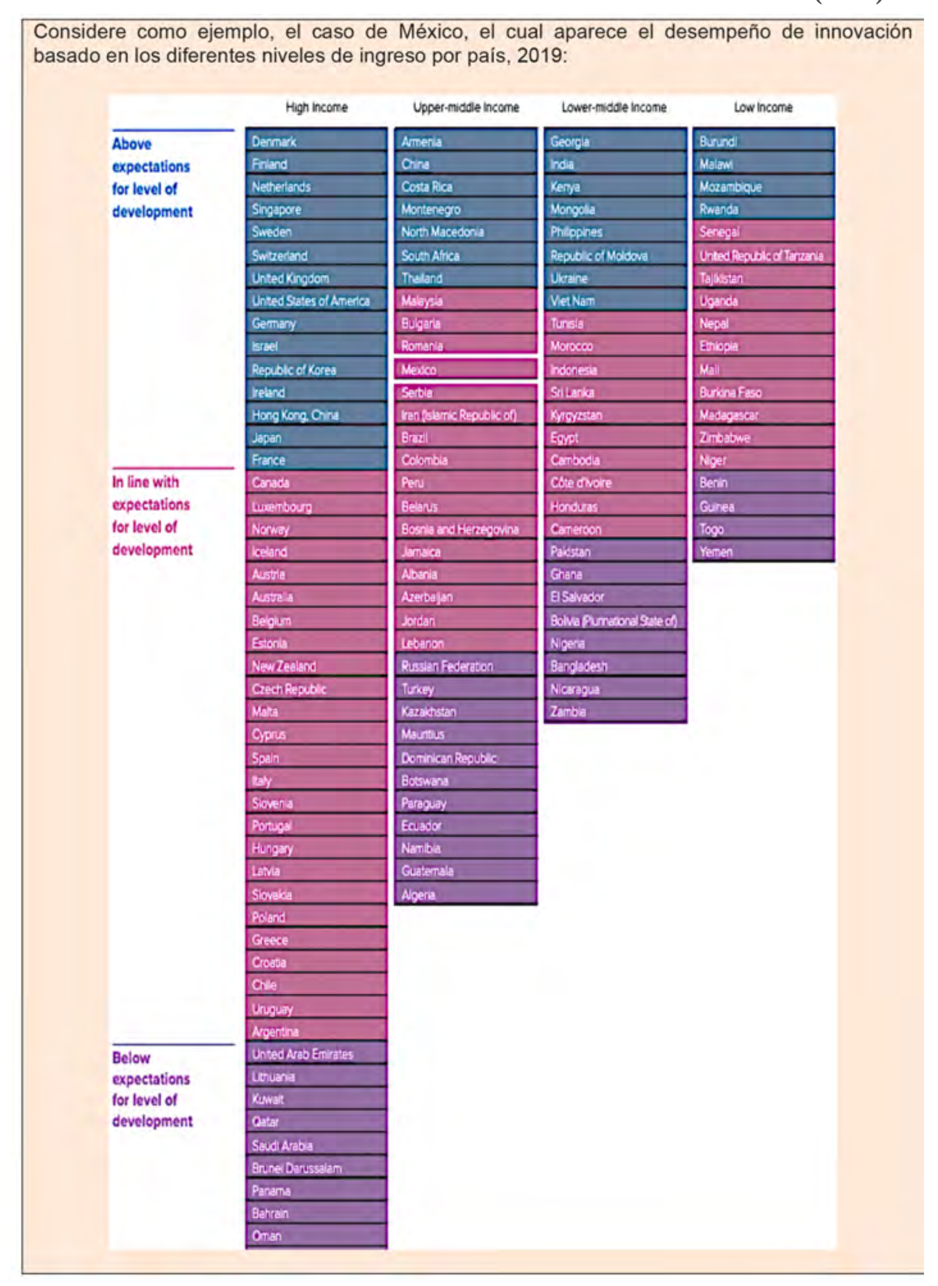
- Por un lado, introduce un mecanismo novedoso para *calibrar* las descripciones verbales y los valores numéricos en *términos de grados de pertenencia a conjuntos* (*set membership scores*) y,
- Por otro lado, permite el uso de valores difusos (*fuzzy*) dentro de los planteamientos analíticos desarrollados para **csQCA**. Esto se consigue equiparando las filas de la *tabla de verdad* a lo que **Ragin (1987)** denomina *espacios de las esquinas vectoriales* (*corner vector spaces*). Si bien cada fila de la *tabla de verdad* genera otra que indica su negación (**A* B * C** lleva a **~ A* B* ~ C**), la *pertenencia difusa a la configuración* **~ (A * B * ~ C**), que puede ser visualmente localizada en el extremo de un vector, indica que *el caso no pertenece a la negación combi-*

*nada de* **(~A)**, *la afirmación de* **B** *y la negación de* **(~ C),** es decir, a **(A* ~ B * C)** que se sitúa en el extremo opuesto.

Como resultado, con conjuntos difusos no se puede hablar de pertenencias nítidas a conjuntos sino que, de forma alternativa, se plantea la proximidad difusa a esquinas que dibujan espacios de pertenencia. Volviendo a la correspondencia de estas esquinas con las *tablas de verdad* formadas por filas, cuando en **fsQCA** se seleccionan casos con la ocurrencia del resultado (o viceversa), se está planteando que estos mismos casos *están más dentro que fuera de la condición* **X** *o más fuera que dentro de la negación de esta misma condición* **(~ X).** Estas dos posibilidades permiten que, en la práctica, **fsQCA** pueda equipararse procedimentalmente al análisis **csQCA.**

El conjunto de *datos difusos*, posee una dualidad simultánea: son *cualitativos y cuantitativos*, ya que incorporan ambos tipos de características en la calibración del grado de pertenencia de las relaciones. Por lo tanto, el conjunto de datos difuso, tiene muchas de las virtudes convencionales de la medición de variables con el tipo de escalas de: intervalo y de razón **(Mejía-Trejo, 2019,** Capítulo 2). Ver **Tabla 4.1.**

**Tabla 4.1. Caso de México de acuerdo al Global Innovation Index (2019)**



Considere como ejemplo, el caso de México, el cual aparece el desempeño de innovación basado en los diferentes niveles de ingreso por país, 2019:

| | High income | Upper-middle income | Lower-middle income | Low income |
|---|---|---|---|---|
| Above expectations for level of development | Denmark | Armenia | Georgia | Burundi |
| | Finland | China | India | Malawi |
| | Netherlands | Costa Rica | Kenya | Mozambique |
| | Singapore | Montenegro | Mongolia | Rwanda |
| | Sweden | North Macedonia | Philippines | Senegal |
| | Switzerland | South Africa | Republic of Moldova | United Republic of Tanzania |
| | United Kingdom | Thailand | Ukraine | Tajikistan |
| | United States of America | Malaysia | Viet Nam | Uganda |
| | Germany | Bulgaria | Tunisia | Nepal |
| | Israel | Romania | Morocco | Ethiopia |
| | Republic of Korea | Mexico | Indonesia | Mali |
| | Ireland | Serbia | Sri Lanka | Burkina Faso |
| | Hong Kong, China | Iran (Islamic Republic of) | Kyrgyzstan | Madagascar |
| | Japan | Brazil | Egypt | Zimbabwe |
| | France | Colombia | Cambodia | Niger |
| In line with expectations for level of development | Canada | Peru | Côte d'Ivoire | Benin |
| | Luxembourg | Belarus | Honduras | Guinea |
| | Norway | Bosnia and Herzegovina | Cameroon | Togo |
| | Iceland | Jamaica | Pakistan | Yemen |
| | Austria | Albania | Ghana | |
| | Australia | Azerbaijan | El Salvador | |
| | Belgium | Jordan | Bolivia (Plurinational State of) | |
| | Estonia | Lebanon | Nigeria | |
| | New Zealand | Russian Federation | Bangladesh | |
| | Czech Republic | Turkey | Nicaragua | |
| | Malta | Kazakhstan | Zambia | |
| | Cyprus | Mauritius | | |
| | Spain | Dominican Republic | | |
| | Italy | Botswana | | |
| | Slovenia | Paraguay | | |
| | Portugal | Ecuador | | |
| | Hungary | Namibia | | |
| | Latvia | Guatemala | | |
| | Slovakia | Algeria | | |
| | Poland | | | |
| | Greece | | | |
| | Croatia | | | |
| | Chile | | | |
| | Uruguay | | | |
| | Argentina | | | |
| Below expectations for level of development | United Arab Emirates | | | |
| | Lithuania | | | |
| | Kuwait | | | |
| | Qatar | | | |
| | Saudi Arabia | | | |
| | Brunei Darussalam | | | |
| | Panama | | | |
| | Bahrain | | | |
| | Oman | | | |

Así, tomando en cuenta el ranking de innovación de México, 2019 como se muestra:

| Country/Economy | Score (0–100) | Rank | Income | Rank | Region | Rank | Median 33.86 |
|---|---|---|---|---|---|---|---|
| Malaysia | 42.68 | 35 | UM | 2 | SEAO | 8 | |
| United Arab Emirates | 42.17 | 36 | HI | 34 | NAWA | 3 | |
| Slovakia | 42.05 | 37 | HI | 35 | EUR | 24 | |
| Lithuania | 41.46 | 38 | HI | 36 | EUR | 25 | |
| Poland | 41.31 | 39 | HI | 37 | EUR | 26 | |
| Bulgaria | 40.35 | 40 | UM | 3 | EUR | 27 | |
| Greece | 38.90 | 41 | HI | 38 | EUR | 28 | |
| Viet Nam | 38.84 | 42 | LM | 1 | SEAO | 9 | |
| Thailand | 38.63 | 43 | UM | 4 | SEAO | 10 | |
| Croatia | 37.82 | 44 | HI | 39 | EUR | 29 | |
| Montenegro | 37.70 | 45 | UM | 5 | EUR | 30 | |
| Russian Federation | 37.62 | 46 | UM | 6 | EUR | 31 | |
| Ukraine | 37.40 | 47 | LM | 2 | EUR | 32 | |
| Georgia | 36.98 | 48 | LM | 3 | NAWA | 4 | |
| Turkey | 36.95 | 49 | UM | 7 | NAWA | 5 | |
| Romania | 36.76 | 50 | UM | 8 | EUR | 33 | |
| Chile | 36.64 | 51 | HI | 40 | LCN | 1 | |
| India | 36.58 | 52 | LM | 4 | CSA | 1 | |
| Mongolia | 36.29 | 53 | LM | 5 | SEAO | 11 | |
| Philippines | 36.18 | 54 | LM | 6 | SEAO | 12 | |
| Costa Rica | 36.13 | 55 | UM | 9 | LCN | 2 | |
| Mexico | 36.06 | 56 | UM | 10 | LCN | 3 | |

Notes: World Bank Income Group Classification (July 2018): LI = low income; LM = lower-middle income; UM = upper-middle income; and HI = high income. Regions are based on the United Nations Classification: EUR = Europe; NAC = Northern America; LCN = Latin America and the Caribbean; CSA = Central and Southern Asia; SEAO = South East Asia, East Asia, and Oceania; NAWA = Northern Africa and Western Asia; SSF = Sub-Saharan Africa.

Se tiene que al recibir una puntuación de relación de 36.06 implica que para el conjunto de países arriba del ingreso de la media (*Upper-middle income*), México se ubica en la línea con expectativas para el nivel de desarrollo (*In line with expectations for level of development*). Sin embargo, aún no es claro si esta relación indica una pertenencia de relación fuerte o débil, aunque es de suponer que para dicho nivel, México se encuentra en una posición no tan débil pero no tan fuerte (ambigua) como para entrar al siguiente nivel de países arriba de las expectativas para el nivel de desarrollo (*Above expectations for level of development*). Este caso, es interesante abordarlo desde el punto de vista del conjunto de *datos difusos*, ya que México tiene que superar los niveles de Rumania, Bulgaria y Malasia y determinar puntuación, sin ambigüedades, que le permita ascender al siguiente nivel.

Fuente: GII (2019).

Una relación de datos difusos, puede verse como una variable continua que ha sido intencionalmente calibrada para indicar el grado de pertenencia de relaciones bien definidas y especificadas. Tal calibración es posible hacerla solo a través del uso de un sólido marco teórico de la investigación y de conocimiento empírico, los cuales son esenciales para la especificación de los tres puntos de corte cualitativos: *completamente perteneciente, completamente no perteneciente y máxima ambigüedad.*

Por ejemplo, casos en los rangos inferiores de una variable continua convencional, pueden estar completamente fuera del conjunto de *datos difusos* en cuestión y con puntuaciones estacionados en **0.0,** mientras que los casos en los rangos superiores de este mismo, la variable continua puede estar completamente en el conjunto de *datos difusos* y puntajes estacionados a **1.0.**

## Ejemplo del conjunto de datos difusos (*fuzzy-sets*)

Para comprender mejor lo antes mencionado, sobre la idea general de lo que representan el conjunto de *datos difusos,* considere los tres valores, que permiten que los casos se encuentren en una *zona de frontera* fuera y dentro del conjunto. Ver **Tabla 4.2.**

Se observa en la **Tabla 4.2**, en lugar de utilizar dos puntajes **0.0** y **0.1**, la lógica de tres valores obliga a un tercer valor proporcional de **0.5** para identificar los casos que no están totalmente dentro o fuera del conjunto en cuestión (compare las *columnas 1 y 2* de la **Tabla 4.2**).

Una modalidad muy apropiada de valores del conjunto de *datos difusos* sets, es la que utiliza cuatro valores numéricos, como se muestra en la *columna 3* de la **Tabla 4.2.** Este esquema utiliza los valores: **0.1, 0.67, 0.33 y 0.0** para indicar: *completamente dentro, más dentro que fuera, más afuera que dentro y completamente fuera,* respectivamente. El esquema de *cuatro valores* es especialmente útil en situaciones donde los investigadores tienen cantidades sustanciales de información acerca de los casos, pero la evidencia se considera que no es sistemática o estrictamente comparable de un caso a otro.

**Tabla 4.2. Conjuntos de datos nítidos (*crisp-sets*) vs. difusas (*fuzzy-sets*)**

| **Columnas** | | | | |
|---|---|---|---|---|
| **Conjunto relaciones nítidas (*crisp-sets*)** | **Conjunto de *datos difusos* (*fuzzy-sets*) de 3 valores** | **Conjunto de *datos difusos* (*fuzzy-sets*) de 4 valores** | **Conjunto de *datos difusos* (*fuzzy-sets*) de 6 valores** | **Conjunto de *datos difusos* (*fuzzy-sets*) *continuo de valores*** |
| 1 = completamente dentro | 1 = completamente dentro | 1 = completamente dentro | 1 = completamente dentro | 1 = completamente dentro |
| — | — | — | 0.8 = principalmente pero no completamente dentro | — Grado en que la pertenencia está más dentro que fuera **0.5** < $X_i$ < 1 |
| — | — | 0.67 = más dentro que fuera | — | |
| — | — | — | 0.6 = más o menos dentro | |
| — | **0.5** = ni completamente dentro ni completamente fuera | — | — | **0.5** = cruce de traslape (*cross-over*) ni dentro ni fuera.. Máxima ambiguedad |
| — | — | — | 0.4 = más o menos fuera | Grado en que la pertenencia está más fuera que dentro **0.0 < $X_i$ < 0.5** — |
| — | — | 0.33 = más fuera que dentro | — | |
| — | — | — | 0.2 = principalmente pero no completamente fuera | |
| 0 = completamente fuera | 0 = completamente fuera | 0 = completamente fuera | 0 = completamente fuera | 0 = completamente fuera |

Fuente: Ragin (2008) con adaptación propia.

Un análisis más detallado, es el que utiliza el conjunto de *datos difusos*, con *seis valores numéricos,* como lo muestra *la columna 4* de la **Tabla 4.2.** Tal y como fue visto en la de *cuatro valores*, la de *seis* utiliza dos estados cualitativos (completamente fuera y completamente dentro). Inserta además, dos estados intermedios:

1. Entre completamente fuera y el punto de cruce de traslape (*cross-over*) (**0.2**= principalmente pero no completamente fuera y **0.4**= más o menos fuera), y
2. Entre el punto de cruce de traslape (*cross-over*) y completamente dentro (**0.6**= más o menos dentro y **0.8**= principalmente pero no completamente dentro).

A primera vista, los conjuntos de *datos difusos* de *cuatro* y *seis* valores, pueden parecer equivalentes a escalas ordinales pero, de hecho, son fundamentalmente diferentes de tales escalas. Una escala ordinal es simplemente un ordenamiento o ranking más de categorías, generalmente sin referencia a criterios, tales como establecer una pertenencia. Al construir *escalas ordinales*, los investigadores no vinculan las categorías al grado de pertenencia en un conjuntos de datos; más bien, las categorías están simplemente ordenadas en relación de uno a otro, dando un orden en el rango.

## Ejemplo constructo de conjunto de datos difusos (*fuzzy-sets*) de seis valores

Por ejemplo, un investigador podría proponer y desarrollar un esquema ordinal de seis niveles de adopción de innovación tecnológica de un país, basadas en el *Technology Readiness Index* o *TIR* **(Parasuraman y Colby, 2015)** y en la curva de difusión de innovación de **Rogers (1962)** y categorizar desde los visionarios innovadores hasta los rezagados. Ver **Tabla 4.3.**

**Tabla 4.3. Propuesta de categorías de adopción de tecnología basada en seis valores de conjunto de datos difusos (*fuzzy-sets*)**

| **Estados de la *Technology Readiness Index* (Parasuraman y Colby, 2015)** | **Estados de la curva de difusión de la innovación de Rogers** | **Estados propuestos en nueva categoría** | **Conjunto de *datos difusos* (*fuzzy-sets*) de 6 valores** |
|---|---|---|---|
| Optimistas | — | Visionarios Innovadores | **1 = completamente dentro** |
| Innovadores | Innovadores | Innovadores | **0.8 = principalmente pero no completamente dentro** |
| | Primeros Adoptadores | Primeros Adoptadores | **0.6 = más o menos dentro** |
| | Mayoría Temprana | Mayoría Temprana | **0.4 = más o menos fuera** |
| Inconformes | Mayoría Tardía | Mayoría Tardía | **0.2 = principalmente pero no completamente fuera** |
| Inseguros | Rezagados | Rezagados | **0 = completamente fuera** |

Fuente: Elaboración propia.

Se destaca la dificultad de la **Tabla 4.3,** particularmente en la *columna* cuatro pudiera traducirse tan directamente como se propone. Sin embargo, sería la base para que el investigador oriente sus esfuerzos a categorizar la adopción de innovaciones tecnológicas en los países emergentes, como México, y que los rangos más bajos de la variable ordinal podrían traducirse completamente fuera del conjunto de países emergentes (puntaje difuso = **0.0**). El siguiente rango podría traducirse a **0.3** en lugar de **0.2**. Los dos primeros puestos podrían traducirse para entrar completamente (puntaje difuso= **1.0**), y así sucesivamente.

En resumen, lo esencial de la traducción de los rangos ordinales a puntuaciones de pertenencia conjuntos de *datos difusos*, depende del ajuste entre el contenido específico de las categorías ordinales y la conceptualización que parte del marco teórico de la investigación, en el etiquetado que proponga para el conjunto de *datos difusos*. Este punto subraya, el hecho

de que los investigadores deben utilizar un marco teórico de la investigación, sólido y sustantivo de conocimiento para calibrar la pertenencia de los conjuntos de *datos difusos*, la cual debe ser útil y reflexivo, nunca mecánico.

Finalmente, el conjunto de *datos difusos* continuo de valores, permite que los casos tomen valores en cualquier lugar en el rango de **0.0** a **1.0**, como se muestra en la última columna de **Tabla 4.2**. El conjunto de *datos difusos* continuo de valores, como todos los conjuntos difusos, utiliza los dos estados cualitativos (totalmente fuera y completamente dentro) y también utiliza el punto de traslape (*crossover*) psara distinguir los casos que están más fuera de los que están más dentro.

## Ejemplo de constructo de conjunto de datos difusos (*fuzzy-sets*) de valores continuos

Considere la pertenencia del conjunto de países desarrollados que realizan las mayores innovaciones y los cuales se basan en el Producto Nacional Bruto per capita (PNB). La traducción de esta variable a puntuaciones del conjunto de *datos difusos*, no se hace ni mecánica ni automáticamente, por lo que se deben observar los siguientes errores a considerar:

**1.** Sería un serio error, por ejemplo, puntear al país menos desarrollado con **(0)** y al más desarrollado con **(1)** y por lo tanto el arreglo con el resto de los países asignar valores de **(0)** a **(1)**, dependiendo de su posición de valores del PNB per cápita.

**2.** Igualmente, sería otro serio error a cometer, basar los puntajes de pertenencia difusa en la orden de rango de los valores PNB per cápita.

Por lo que deberá a abocarse a realizar la puntuación de al menos tres cortes cualitativos basados en la distribución del PNB per cápita. Vea la **Tabla 4.4.**

**Tabla 4.4. Pasos para llevar a cabo un correcto planteamiento de conjunto de datos difusos (*fuzzy-sets*) por puntos de corte cualitativos**

| Paso | Descripción |
|---|---|
| **1** | En el cual se alcanza la pertenencia completa del grupo (basado en un país claramente desarrollado) |
| **2** | En el cual se alcanza la no pertenencia completa del grupo (basado en un país claramente no desarrollado) |
| **3** | En el cual se alcanza la *máxima ambigüedad del grupo*, considerando si un país está más adentro o más afuera del conjunto países desarrollados (puntajes de pertenencia de **0.5**, como punto cruce de traslape (*cross-over*)) |

Fuente: Ragin (2008) con adaptación propia.

Al hacer específicos los puntos de corte cualitativos, se permitirá hacer distinción entre una variación *relevante* de una *irrelevante*. La variación en el PNB per cápita entre los países desarrollados (sin ambigüedad) no será relevante para graduar al conjunto de dichos países, al menos, desde el punto de vista del conjunto de *datos difusos*; si un país es claramente desarrollado, se le aplica su pertenencia como tal. Del mismo modo, la variación en el PNB per cápita entre los países claramente no desarrollados, también es considerada no relevante en el conjunto de los países desarrollados. Por lo tanto, la investigación que utilice el conjunto de *datos difusos*, no le será suficiente desarrollar simplemente escalas que muestren las posiciones que se relacionen entre sí. Será necesario utilizar los conocimientos basados en el marco teórico de la investigación, así como empírico para mapear los enlaces entre puntajes específicos de las variables continuas (por ejemplo, el *Global Innovation index*, o **GIN, 2019**) y puntuaciones específicas de pertenencia del conjunto de *datos difusos*, por ejemplo, pertenencia completa en el conjunto de países desarrollados; se deduce que, cuando un investigador reconceptualiza y vuelve a etiquetar un conjunto (por ejemplo, cambiando el enfoque del conjunto de países desarrollados al conjunto de países de emergentes), los puntajes de pertenencia cambian en consecuencia, aunque la variable (por ejemplo, PNB per cápita) sea la misma.

## Fundamentos y operaciones del conjunto de datos difusos (*fuzzy-sets*)

Cuando se usan conjuntos de *datos difusos*, para evaluar las relaciones basadas en marco teórico de la investigación, ambos, los resultados y las *condiciones causales* pueden ser representadas en términos de puntuaciones de pertenencia (los conjuntos de datos nítidos, pueden ser incluidos con las condiciones de conjuntos de *datos difusos*, en un análisis difuso). Considere por ejemplo, las primeras cinco *columnas* de la **Tabla 4.5.** que corresponde al archivo de trabajo: **empresas innovadoras cluster.cvs.**

En la **Tabla 4.5,** se muestra una matriz de conjunto de *datos difusos*, para el caso hipotético de cómo un grupo de empresas innovadoras ubicadas dentro de un clúster de innovación tecnológico de manufactura, presenta su toma de decisiones dentro de un clúster de innovación, siendo de principal interés, el grado de pertenencia en el conjunto de empresas innovadoras con diversos niveles de *calidad de votación por liderazgo tecnológico,* **W.** Este análisis se centra en las condiciones ligadas a las votaciones de las empresas innovadoras clúster con bajo nivel de calidad de votación (solo una empresa es = **1.0**). Las *condiciones causales* usadas en el ejemplo, son:

- **a.** Grado de pertenencia en el conjunto de empreas innovadoras clúster con sólidas fortalezas en resultados de sus consejos de administración, **U.**
- **b.** Grado de pertenencia en el conjunto empresas innovadoras clúster con empleo de un alto grado tecnológico de manufactura, **M.**
- **c.** Grado de pertenencia en el conjunto de empresas innovadoras clúster por su nivel de influencia, **A.**
- **d.** Grado de pertenencia en el conjunto de empresas innovadoras clúster con diferencia en sus ingresos por innovación, **I.**

**Tabla 4.5. Datos de conjunto de datos difusos (*fuzzy-sets*) para el caso de decisión de clúster de innovación**

| Empresas innovadoras Clúster | Columnas | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | Calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado de tecnología manufactura | Fortaleza resultados consejo de administración | (~ M) | (A) * ( ~M) | (A) + (~M) |
| | W | A | I | M | U | | | |
| **A** | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 | 0.6 | 0.6 | 0.8 |
| **B** | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 | 0.8 | 0.6 | 0.8 |
| **C** | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 | 0.8 | 0.6 | 0.8 |
| **D** | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 | 0.8 | 0.6 | 0.8 |
| **E** | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 | 0.6 | 0.6 | 0.6 |
| **F** | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 | 0.2 | 0.2 | 0.2 |
| **G** | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 | 0.8 | 0.4 | 0.8 |
| **H** | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 | 0.8 | 0.6 | 0.8 |
| **I** | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 | 0.4 | 0.4 | 0.6 |
| **J** | 0.0 | 0.8 | 0.4 | 0.8 | 1.0 | 0.2 | 0.2 | 0.8 |
| **K** | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 | 0.2 | 0.2 | 0.6 |
| **L** | 1.0 | 1.0 | 0.8 | 0.4 | 0.2 | 0.6 | 0.6 | 1.0 |

Fuente: Elaboración propia.

En el conjunto de empresas innovadoras clúster, con sólidas *fortalezas en resultados de sus consejos de administración* **U** y con empleo de un *alto grado tecnológico de manufactura* **M,** tienden a tener una *fuerte calidad de votación por liderazgo tecnológico* **W,** mientras las que tienen *nivel de influenci*a **A** y *diferencia de ingresos por innovación* **I,** tienden a socavarla. Todos los conjuntos de *datos difusos* en este análisis, son de seis valores y están basados en características generales de las empresas. Mientras más detalle sean las graduaciones son posibles con estos datos (como la *columna 4* de *valores contínuos*, de la **Tabla 4.2.**), el intento aquí es solo académico para demostrar las operaciones con los conjuntos de *datos difusos* a partir de un conjunto de datos. La presentación no se centra por el momento en cómo se calibrar el conjunto de *datos difusos* o en el problema de cuales *condiciones causales* podrían proporcionar la mejor especificación de las columnas de la **Tabla 4.5**. Tres operaciones comunes se realizan en conjuntos de *datos difusos*, los cuales son:

**a.** Conjuntos de relación de negación,
**b.** Conjuntos de relación de intersección (lógico **AND**) y,
**c.** Conjuntos de relación de unión (lógico **OR**).

La discusión de estas tres operaciones brinda importantes conocimientos previos para entender cómo trabajar con los conjuntos de *datos difusos*.

## Negación

Al igual que los conjuntos de datos nítidos convencionales, los conjuntos de *datos difusos*, se pueden negar. Con los conjuntos de datos nítidos, la negación cambia los puntajes de pertenencia de **1.0 a 0.0** y de **0.0 a 1.0**. Por ejemplo, la negación en los conjuntos de datos nítidos, de las empresas innovadoras en clúster es las no empresas innovadoras clúster. Este simple principio matemático también se aplica al álgebra difusa, pero el valor numérico relevante no está restringido a los dos valores booleanos **0.0**

**y 1.0**; más bien, se extienden a valores entre **0.0 y 1.0**. Para calcular la pertenencia de un caso en la negación del conjunto de *datos difusos* de **M,** simplemente reste su pertenencia en el conjunto **M** de **1.0**, como sigue:

Pertenencia en el conjunto
**(~ M)** = **1.0** – (pertenencia en el conjunto **M** ), o **~ M = 1.0 – (M)**

Donde ~ es el signo de negación. Así, por ejemplo, la *empresa innovadora del clúster* **L,** con puntaje de **0.4** en el conjunto *de empresas innovadoras con grado tecnológico de manufactura* **M** alto, por lo tanto, tiene un puntaje de **0.6** en el conjunto de **M** no alto. Otro ejemplo sería, el examinar los datos de la sexta columna de la **Tabla 4.5** el cual muestra el puntaje de la pertenencia negada del conjunto de *empresas innovadoras con grado tecnológico de manufactura* **M** alto para todas las doce empresas innovadoras del clúster. Vea la *columna 6*, donde la negación del conjunto, es etiquetado como **(~ M),** para del conjunto de **M** no alto. Así, por ejemplo, en el conjunto de *datos difusos*, su pertenencia en el conjunto de empresas por ingresos de innovación es **0.91** y es de **0.09** en el conjunto de empresas que no tiene ingresos de innovación.

## Conjunción AND lógico

Los conjuntos compuestos se forman cuando dos o más conjuntos se combinan, una operación comúnmente conocida como conjunto de intersección. Un investigador interesado en el destino de la calidad de votación es relativamente raro. Otra configuración requerida puede ser el elaborar una lista de países que combinen el empleo de casos *de empresas innovadoras con grado tecnológico de manufactura* **M,** no alto y *nivel de influencia* **A** alto. Convencionalmente, estas empresas serían identificados utilizando conjuntos de datos nítidos mediante la tabulación cruzada de las dos dicotomías (*grado tecnológico de manufactura* **M** no altas vs. alta y *el nivel influencia* **A** alta vs. no alta, y ver qué empresas innovadoras del clúster se

encuentran en *grado tecnológico de manufactura* **M** no alto con el *nivel de influencia* **A** alta. Esta celda, mostraría los casos que existen en la intersección de las dos conjuntos de datos nítidos. Con el conjunto de *datos difusos*, la **AND** lógica es lograda al tomar el mínimo puntaje de pertenencia de cada caso en los conjuntos a ser combinados. Por ejemplo, si la pertenencia de una empresa en el conjunto de empresas innovadoras del clúster con un *grado tecnológico de manufactura* **M** no alto es de nivel **0.6** y su pertenencia en dicho grupo de empresas con nivel de influencia **A** alto es **1.0**, entonces su pertenencia en el conjunto de empresas innovadoras del clúster que combinan estos dos rasgos es el *menor* de estos dos puntajes, es decir **0.6.** Un puntaje de **0.6** indica que este caso, está todavía más dentro que fuera de la intersección. Para ilustrar este principio, considere la *columna 7* de la **Tabla 4.5.** el cual muestra la operación lógica de conjunción **AND**. Esta columna muestra, que la intersección del puntaje de la pertenencia negada del conjunto de *empresas innovadoras del clúster con grado tecnológico de manufactura* **(~M)** no alto con el conjunto de empresas innovadoras del clúster con *influencia* **A** alta, producen pertenencias en el conjunto de empresas que combinan estos dos rasgos. La expresión algebraica para esta intersección es:

**(A * ~M)**

El asterisco **(*)** se usa para indicar el conjunto como intersección (combinación de aspectos). En los conjuntos de *datos difusos*, el **AND** lógico, se logra tomando el *mínimo* de puntuación de pertenencia de cada caso, en los conjuntos que se intersectan. Por ejemplo, si la pertenencia de una empresa innovadora pyme en el conjunto de empresas de alta tecnología es **0.34** y su pertenencia en el conjunto de empresas por ingresos de innovación es **0.91**, su pertenencia en el conjunto de empresas de alta tecnología y por ingresos de innovación es la *menor* de estas dos puntuaciones, o sea **0.34.**

## Unión OR lógico

Dos o más conjuntos también se pueden unir mediante la lógica **OR**, la unión de conjuntos. Por ejemplo, un investigador podría estar interesado ya sea del conjunto de *empresas innovadoras del clúster con grado tecnológico de manufactura* **(~M)** no alto o del conjunto de *empresas innovadoras del clúster con influencia* **A** alta, según la conjetura que se haga de estas dos condiciones, que pueden ofrecer bases equivalentes y sustituibles para algún resultado (por ejemplo, *votación de calidad débil,* **W**). Cuando se usan en conjuntos de *datos difusos*, el **OR** lógico, este dirige la atención del investigador al máximo de cada caso de pertenencias en los conjuntos de componentes. Es decir, la pertenencia de un caso en el conjunto formado por la unión de dos o más conjuntos de *datos difusos*, es el máximo valor de sus pertenencias en los conjuntos de componentes. Por lo tanto, si una empresa innovadora del clúster tiene un puntaje de **0.2** en el conjunto de empresas innovadoras con influencia **A** y un puntaje de **0.8** en el conjunto de empresas innovadoras del clúster con grado tecnológico de manufactura **M** no alto, tiene una puntuación de **0.8** en el conjunto de empresas innovadoras del clúster que tienen cualquiera de estos dos rasgos. Por ejemplo, del uso de la lógica **OR**, considere la *columna 8* de datos de la **Tabla 4.5.** Esta columna muestra, a las empresas innovadoras del clúster que tienen o un porcentaje de *grado tecnológico de manufactura* **M** no alto o un *nivel de influencia* **A** alta. La expresión algebraica es:

**(~M) + (A)**

Donde el signo **(+)** es utilizado para indicar el **OR** lógico. Con conjuntos de *datos difusos*, el investigador se enfoca en el máximo de pertenencias de cada caso, de los componentes. Es decir, la pertenencia en el conjunto formado por la unión de dos o más conjuntos de componentes es el valor máximo de las pertenencias del caso en los conjuntos de componentes. Por lo tanto, si la pertenencia de una empresa innovadora pyme en el conjunto de empresas de alta tecnología es **0.34** y su pertenencia en el conjunto de empresas por ingresos de innovación es **0.91**, su pertenencia en el conjunto de empresas de alta tecnología y por ingresos de innovación es la *mayor* de estas dos puntuaciones, o sea **0.91**.

## Calibración fsQCA de las condiciones

**Ragin (1987)** ofrece una manera de operacionalizar la asignación de valores difusos siguiendo la idea de *grados de pertenencia* a los conjuntos. Lo que hace con esta propuesta, es simplificar el uso de los *términos lingüísticos* empleados en otras disciplinas, para permitir su *calibración* en valores difusos. En términos lingüísticos esta *vaguedad* se puede presentar, en torno a tres dimensiones o cuantificadores vagos. Algunos se presentan en forma de cuantificación: todos, muchos, la mayoría, la mitad, unos pocos, etcétera. Otros se refieren a la frecuencia, a saber, siempre, frecuentemente, a menudo, ocasionalmente, raramente, nunca. Por último, otro tipo engloba posibilidades tales como probablemente, improbable, cierto, incierto, seguro, inseguro, etcétera.

Calibrar no es más que transformar una variable de intervalos o de razón, sin aparentes límites, en una variable que oscile entre **(0,1)**. Para ello es necesario distinguir entre variaciones *relevantes e irrelevantes* en el momento de definir la pertenencia de los casos a los conjuntos (*set membership scores*), es decir, es preciso indicar en la variable original dónde se sitúa teóricamente la pertenencia total al conjunto que previamente se ha definido, dónde se encuentra el punto que marca la total exclusión del conjunto y dónde localizamos la zona intermedia en la que no está clara la pertenencia o exclusión de dicho conjunto. Para ello, al tratarse de un procedimiento teórico, el investigador debe aportar sentido (y contextualización) para definir los llamados *anclajes o referencias teóricos* a una matriz de datos formada por índices o datos numéricos. Calibrar es una operación específica para **fsQCA** que no encontramos en **csQCA**: los *anclajes o referencias teóricos* en *conjuntos nítidos* (*crisp sets*) se encuentran colapsados en uno cuando se produce la diferenciación entre **(0)** y **(1)**. No obstante, como argumenta **Rihoux (2008, p. 731)**, la calibración de valores difusos no es mecánica; es preciso apoyarse de *conocimiento teórico* y *empírico sustantivo* para definir dónde se sitúan los *anclajes o referencias teóricos*.

Es importante tener en cuenta que *un valor difuso no es un variable continua o numérica,* pues **fsQCA** opera dentro de los métodos de conjuntos.

**Ragin (2000, p. 8)** entiende que los *conjuntos difusos* combinan valoraciones cuantitativas y cualitativas durante su construcción pues, de una parte, estos requieren la fijación cualitativa de tres valores de referencia (completa pertenencia /punto de máxima indefinición/ completa exclusión), así como variación cuantitativa entre ellos. Esto lleva a diferenciar los casos por su tipología (diferencia cualitativa) y por su grado (diferencia cuantitativa) entre casos cualitativamente idénticos **(Schneider y Wagemann, 2012, p. 27)**.

Siempre va a ser necesario indicar en qué momento el caso estará totalmente dentro de **A,** cuándo estará totalmente fuera de **(A) (~ a, a)** y cuándo se encuentra en una posición *intermedia* de completa borrosidad (*fuzziness*). La simple presentación de datos numéricos o de códigos cualitativos (*linguistic qualifiers*) no se traduce en valores *difusos*, por lo que el análisis resultará erróneo. Se entiende por *códigos cualitativos,* la traducción de *variables categóricas* a *puntuaciones interválicas* siguiendo un cierto orden de *pertenencia a conjuntos dicotómicos*. Estos códigos cualitativos deben ser calibrados, es decir, el investigador tiene que incluir los anclajes o referencias teóricos para convertirlos en conjuntos de datos *difusos. La justificación de dónde los situamos debe estar detallada en la discusión metodológica de toda investigación* **fsQCA**. Esta justificación es necesaria, pues no existe un código al que recurrir en el momento de calibrar datos. Por ejemplo, no se encuentra disponible una guía en ningún manual de referencia que indique con exactitud en qué valor concreto se puede establecer la frontera entre una empresa caracterizada por *actividad de innovación moderada y otro con una actividad de innovación elevada.* Se desconoce también, a partir de qué porcentaje se considera que una empresa realizan actividades de innovación cerrada y en qué momento pasa a ser innovación abierta.

Un recurso fácil es fijar la *zona de borrosidad* justo a la mitad entre los dos valores extremos con la *media estadística*, o alternativamente la *mediana*, pero esta estrategia *puede no resultar ser la más adecuada*, pues todo depende del objetivo empírico de la condición. Cabe mencionar, que tampoco está justificado el uso de la herramienta *Thresholdsetter* habili-

tada en **TOSMANA** para calibrar valores difusos. Dicha herramienta está pensada para análisis **csQCA** y **mvQCA**. Si se toma, por ejemplo, el gasto que realizan los centros públicos de investigación de un país, el establecimiento de las referencias teóricas variará dependiendo de si se quiere capturar qué centros públicos son los que menos gastan, interés que lleva a plantear una localización de casos poco poblada cuanto más cerca de **(1)** y mucho más poblada al acercarse a **(0)** con un *cross-over point* intencionadamente desplazado hacia cifras de gasto público bajas. De lo contrario, si el interés es construir una condición que relativice el gasto público por centro público de investigación entre regiones del país es decir, sin una voluntad de enfatizar clústeres extremos dentro de la lista de regiones, entonces es acertado situar el *cross-over point* en algún lugar intermedio. A menudo son los propios datos los que ofrecen variaciones significativas (*saltos*, *lagunas*) que ayudan a colocar los *anclajes o valores de referencia teóricos*.

De esta forma, el proceso de calibración acepta matrices de datos fruto de investigaciones *cualitativas* e investigaciones *cuantitativas*. Resultado de investigaciones *cualitativas* las *matrices de datos* estarán conformadas por *códigos cualitativos* (también llamados *clasificaciones cualitativas*) que el investigador habrá creado para clasificar los casos. Estos códigos responden a variaciones conceptuales. Asimismo, acostumbran a presentarse de manera interválica capturando cierta graduación entre elementos. Al ser definidos por el investigador, estos códigos pueden estar escalados de **(0)** a **(1)**, pero no es del todo necesario si el objeto de estudio no lo recomienda. Una vez definidos, deberán ser revisados teóricamente para corresponderse con valores difusos. Entonces sí que será necesario proponer un *rango de pertenencia*.

Por otro lado, es posible crear *matrices de datos* a partir de valores numéricos. La mayoría de los datos demográficos y económicos son de este tipo. Por ejemplo, datos sobre población, niveles de riqueza, ingresos de las familias, capacidad de ahorro familiar, gasto en inversiones públicas, niveles de endeudamiento público, número de empresas en un país, número de población activa, número de nacimientos, esperanza de vida, etcétera.

En innovación por ejemplo, el Global Innovation Index se basa en datos previos de: ambiente político, ambiente regulatorio, ambiente de negocios, educación, investigación y desarrollo, tecnologías de información y comunicaciones, infraestructura, sustentabilidad ecológica, crédito, inversión, competitividad del mercado, trabajadores del conocimiento, vínculos de innovación y absorción del conocimiento entre otros.

Si bien el investigador *no requiere crear códigos cualitativos* previos en el momento de conseguir estos datos, sí será necesario un esfuerzo para *identificar variaciones relevantes* dentro de los datos. Para esto se necesita definir con exactitud el *sentido positivo o negativo de la calibración*, es decir, plantear *si la calibración requiere un punto extremo positivo* (*positive end-point*) *o uno negativo* (*negative end-point*). En la práctica, con unos datos que midan por ejemplo el grado de innovación de diferentes países, el investigador tendrá que decidir si prefiere *calibrar de forma positiva los países* (de menor a mayor grado de innovación) o si, de lo contrario, su condición captura grados de pertenencia a países con un grado de innovación bajo. Esto supone calibrar inversamente: los países con mayor grado de innovación estarán más próximos al valor **(0)** mientras que los países con un bajo grado de innovación estarán más próximos al valor **(1).** En cualquier caso, *el investigador tendrá que considerar cómo se define el extremo de referencia para una buena lectura de los valores a lo largo del análisis.*

## Modelos de calibración

La definición y asignación de valores *difusos* puede tomar varias formas **(Ragin, 2000, 2008, 2009b)**. Si bien los valores *dicotómicos* oscilan entre **(0)** y **(1),** es posible *operacionalizar los grados de pertenencia* en *conjuntos difusos* dependiendo de las necesidades del investigador y la singularidad de los datos.

La forma más sencilla de asignar *grados de pertenencia* (o *escalas difusas*) responde al modelo *conjunto difuso de tres valores* (*three-value fuzzy set*) que establece tres puntos de referencia para cuando los datos están totalmente dentro del conjunto **(1)**, cuando se encuentran en un estado

de completa borrosidad **(0.5)** y cuando se considera que están totalmente fuera del conjunto **(0)**. Aunque este modelo puede parecer el más sencillo, **Schneider y Wagemann (2012, p. 28)** *advierten del riesgo de asignar un valor fuzzy de* **(0.5)** *a los casos, pues se les estaría adjudicando un estatus cualitativo indefinido* y, por lo tanto, por un lado, se les *declara débiles conceptualmente* y, por otro lado, pueden presentarse algunos problemas en el momento de analizar las *tablas de verdad*. Cabe decir que los anclajes o *valores de referencia extremos* **(0)** *y* **(1)** *no tienen que corresponder con el valor máximo y mínimo en la variable original como tampoco a los extremos de una hipotética clasificación cualitativa.* La definición de tres *valores de referencia* son orientaciones a partir de las cuales relacionar los casos más semejantes. De otro lado, como ya se ha mencionado, la necesidad de encontrar un *punto intermedio* que funcione como *cross-overpoint* (*o punto de traslape*) precisa de un argumento teórico y *no basta con recurrir por defecto a medidas estocásticas como la media o la mediana.*

Para una calibración algo más refinada y menos restrictiva, existe la posibilidad de asignar una escala de *cuatro valores* (*four-value fuzzy set*) de la siguiente manera (Ver capítulo 6): **(1)** corresponde a la pertenencia total al conjunto; **(0.67)** corresponde a casos que están más dentro que fuera; **(0.33)** se reserva para casos que están más fuera que dentro del conjunto; y **(0)** se asigna a casos que están totalmente fuera del conjunto. Nótese que en esta propuesta de calibración, la zona de mayor borrosidad no está delimitada específicamente en el **(0.5)**; esta se enmarca en el *área indeterminada* entre **(0.67)** y **(0.33)**. Con ello los *casos intermedios* se localizan en *posiciones de mayor o menor pertenencia en el conjunto, con lo que, definitivamente, no serán tratados como casos completamente difusos.*

En algunas situaciones, la *matriz de datos* es algo más compleja y es necesario contar con una *calibración con más de cuatro valores*. Para ello los investigadores pueden fijar *conjuntos difusos de seis valores* (*six-value fuzzy set*) en cuya modalidad estándar propone la siguiente graduación: **(1)** corresponde a total pertenencia; **(0.9)** se asigna a casos que están mayoritariamente pero no totalmente dentro del conjunto; **(0.6)** sirve para fijar una pertenencia más o menos dentro del conjunto, como también se recurre al

valor **(0.4)** para casos en una situación intermedia; el valor **(0.1)** se asigna a casos que están casi fuera del conjunto; y **(0)** se reserva para casos totalmente fuera.

Finalmente, los conjuntos difusos continuos permiten que los casos tomen valores en cualquier punto a lo largo del intervalo **(0,1)**. Al no ser una asignación restrictiva, el investigador establece el criterio de total pertenencia en el valor **(1)**, el punto intermedio se fija en **(0.5)** y el valor de total exclusión en **(0)**. *Este requisito puede hacer pensar que la calibración continua es similar al conjunto difuso de tres valores*; *sin embargo, la diferencia radica en el hecho de que en la anterior forma de calibración los valores* **(1)**, **(0.5)** y **(0)** *funcionaban como contenedores conceptuales relativamente amplios para organizar los casos a partir de ellos*. En esta nueva propuesta, de lo contrario, los valores de referencia teóricos establecen tres puntos que sirven como *referencias de pertenencia* con un significado conceptual propio para después relacionar los casos entre sí. Ello se consigue estableciendo dos grandes áreas de borrosidad en las que, por un lado, los casos pueden estar gradualmente *más dentro que fuera* **($0.5 < X_i < 1$)** y, por otro lado, los casos pueden graduarse dentro del área *más fuera que dentro* del conjunto **($0 < X_i < 0.5$).**

Cabe decir que *se pueden emplear diferentes modelos de calibración para cada una de las condiciones*, si se desea y está justificado. Se puede encontrar una condición calibrada en seis valores y otra *fuzzificada* en torno a tres valores. También es posible introducir *valores dicotómicos*. En cualquier caso, el investigador debe ser consciente del hecho de que la *calibración* puede terminar afectando los análisis y cómo los casos se relacionan entre ellos. Por este motivo la elección de la estrategia de *calibración* no puede ser un resultado de una decisión incongruente. El investigador debe inspeccionar las consecuencias de la *calibración* para mejorarla si es preciso. La relación entre teoría y práctica, o el *ir y volver* de forma constante, adquiere total significado en este punto.

## Métodos de calibración

Tras la decisión sobre el modelo de calibración, el investigador debe elegir el mejor método para transformar variables originales en condiciones difusas. El modelo de calibración será útil, entonces, para saber cuántos valores son necesarios, y el método será empleado para calibrar la variable original en una condición difusa de tres, cuatro, seis o continuos valores. Como ahora se explica, la calibración puede realizarse de forma teórica o semiautomática, con ayuda o no del software habilitado para ello **(fsQCA 3.0, STATA, R).** Todos los métodos de calibración son buenos y bienvenidos si su aplicación está justificada. La sencillez o la falta de dominio no debe ser una justificación válida, como tampoco hay un método superior a otro y, por consiguiente, con mayor aceptación académica. El investigador debe recurrir al mejor método de calibración: ese método que se adapte mejor, en un sentido teórico y práctico, al modelo de calibración. Cuando uno y otro están perfectamente justificados, entonces se tendrá la certeza de no haber errado.

Por lo general, la literatura **(Ragin, 2008; Schneider y Wagemann, 2012)** reconocen tres métodos de calibración:

**1.** *El método de calibración cualitativa, o método directo de atribución,* se fundamenta en el conocimiento y la experiencia del investigador para designar la localización de los *anclajes o los valores de referencia teóricos* **(Verkuilen, 2005)**. Aunque resulte tentador establecer valores de referencia teóricos de forma aleatoria, el investigador está obligado a dotar de *sentido teórico* a cada uno de dichos valores, bien porque estos se adapten a una previa clasificación conceptual elaborada por el investigador en la fase de diseño de la investigación o bien porque existan separaciones evidentes entre los datos. El *método de calibración cualitativa* permite que el investigador, siguiendo argumentos teóricos y experiencia empírica, decida asignar pertenencias a valores difusos. El *método cualitativo* se utiliza cuando no existe una previa medición cuantitativa de los casos. Si hay una medición cuantitativa susceptible de ser calibrada, lo más adecuado es usar el método directo o el indi-

recto. El *método cualitativo* se puede usar cuando se está ante un fenómeno cualitativo que puede ser entendido de forma ordinal, pero para el cual no disponemos de una medición cuantitativa. En ese caso, el investigador debe atribuir grados de pertenencia basándose en su conocimiento cualitativo de los casos, siguiendo un criterio lógico y justificado de atribución de valores. De nuevo hay que recordar que este método, a pesar de su sencillez procedimental, *es el que más necesita de argumentación por parte del investigador*, precisamente porque la calibración es producto de su conocimiento y experiencia. El investigador debe convencer del sentido cualitativo que tiene cada uno de los grupos que ha creado; es importante evitar críticas sobre la falta de precisión o la manipulación intencionada.

**2.** *El método directo de calibración,* emplea una función logística para encuadrar los datos dentro de *tres valores de referencia teóricos* establecidos entre **(1)** para total pertenencia, **(0.5)** punto de indiferencia y **(0)** total exclusión. Este método puede realizarse con la herramienta **Compute** del software **fsQCA 3.0**, que es quizás la manera más sencilla de utilizar. Debido a la aplicación de la *función logística*, los valores de referencia se sitúan **en 0.95, 0.5 y 0.05**. Como paso intermedio, usa estimaciones del *logaritmo de probabilidad de pertenencia total en un conjunto* **(Ragin, 2008, p. 87)**. De forma sencilla, el primer paso es traducir *categorías verbales en grados de pertenencia a conjuntos*; se tiene que establecer que el *umbral de total pertenencia* (*categoría verbal*) se traduce en **0.95** *grados de pertenencia* y así de forma sucesiva con el resto de las *etiquetas verbales.* En el momento de proceder con la calibración, este paso ya deber estar definido *al seleccionar el modelo de calibración* más apropiado para la matriz de datos. A continuación se obtienen los *posibilidad de pertenencia (odds of membership)* de cada uno de los grados de pertenencia. Esto se consigue de forma muy sencilla aplicando la siguiente fórmula **(Ragin, 2008, p. 87)**:

**posibilidad de pertenencia *(Odds of membership)* =**
**(grado de pertenencia) / (1-(grado de pertenencia)]**

Para un grado de pertenencia de **(0.95)** se obtiene una *posibilidad de pertenencia (odd of membership)* de **19**, para un grado de **(0.67)** se obtiene **2.03**, para un *cross-over point* de **(0.5)** se obtiene **1**, para un grado de **(0.33)** se obtiene **0.49** y para un grado de **(0.05)** se obtiene **0.05**. La categoría verbal *total pertenencia* obtiene una *posibilidad de pertenencia (odd of membership)* de **100**, mientras que la *total exclusión* recibe **0.01**. Por último, se calculan los *logaritmos naturales de esta posibilidad de pertenencia (odds of membership)*. Estos *logaritmos naturales* serán útiles en pasos posteriores. La **Tabla 4.6** resume estas operaciones de transformación matemática de siete etiquetas verbales.

**Tabla 4.6. Transformaciones matemáticas de etiquetas verbales**

| Categorías verbales | Grados de pertenencia | Posibilidad de pertenencia *(odds of membership)* | Logaritmo natural posibilidad de pertenencia *(odds of membership)* |
|---|---|---|---|
| **Total pertenencia** | 0.99 | 100 | 4.605 |
| **Umbral de total pertenencia** | 0.95 | 19 | 2.944 |
| **Más dentro que fuera** | 0.67 | 2.03 | 0.708 |
| **Total indefinición** | 0.5 | 1.00 | 0 |
| **Más fuera que dentro** | 0.33 | 0.49 | -0.713 |
| **Umbral de total exclusión** | 0.05 | 0.05 | -2.995 |
| **Total exclusión** | 0.01 | 0.01 | -4.605 |

Fuente: Medina (et al., 2017).

Siguiendo con el ejemplo, sirva la *matriz de datos hipotético* de la **Tabla 4.7.**

**Tabla 4.7. Ejemplo hipotético de configuración contradictoria**

| Productos de innovación de empresa ejemplo anual | Columna | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | Ingresos anuales por innovación (musd) | Desviaciones del punto de cruce (cross-over point) | Escalares Punto cruce = 108 superior = 150; Inferior= 80 | Producto Posibilidad de pertenencia por logaritmo natural | Grado de pertenencia |
| | (I) | D = ( I – 5000 ) | (X) Límite superior Límite=150 inferior= 80 | P = (D) * (X) | Y = [exp (P) ] / [ 1+ exp (P)] |
| **1** | 189 | 81 | 0.0712 | 5.7672 | 0.9968 |
| **2** | 154 | 46 | 0.0712 | 3.2752 | 0.9636 |
| **3** | 152 | 44 | 0.0712 | 3.1328 | 0.9582 |
| **4** | 132 | 24 | 0.0712 | 1.7088 | 0.8467 |
| **5** | 126 | 18 | 0.0712 | 1.2816 | 0.7827 |
| **6** | 112 | 4 | 0.0712 | 0.2848 | 0.5707 |
| **7** | 112 | 4 | 0.0712 | 0.2848 | 0.5707 |
| **8** | 111 | 3 | 0.0712 | 0.2136 | 0.5532 |
| **9** | 110 | 2 | 0.0712 | 0.1424 | 0.5355 |
| **10** | 107 | -1 | 0.1069 | -0.1069 | 0.4733 |
| **11** | 106 | -2 | 0.1069 | -0.2138 | 0.4468 |
| **12** | 105 | -3 | 0.1069 | -0.3207 | 0.4205 |
| **13** | 99 | -9 | 0.1069 | -0.9621 | 0.2765 |
| **14** | 95 | -13 | 0.1069 | -1.3897 | 0.1995 |
| **15** | 93 | -15 | 0.1069 | -1.6035 | 0.1675 |
| **16** | 76 | -32 | 0.1069 | -3.4208 | 0.0317 |
| **17** | 60 | -48 | 0.1069 | -5.1312 | 0.0059 |

Fuente: Elaboración propia.

Por lo que se requiere:

**a.** Lo primero es definir la localización del punto de traslape (*cross-over point*) en la *matriz de datos*. Esto servirá para calcular las desvia-

ciones de cada caso del *punto de cruce* (*cross-over point*). A efectos de ejemplo este se establece en **108** *ingresos anuales por innovación.*

**b.** Seguidamente se calcula la desviación que hay entre el *punto de cruce (cross-over point)* y los valores que se han asignado como *umbrales de total pertenencia y exclusión*. Claro está, estos umbrales no siempre tienen que ser los valores extremos en la matriz de datos sino aquellos que muestren en una variación relevante. Se seleccionan **150** *ingresos anuales por innovación* como *umbral de total pertenencia como límite superior*, que es una desviación de **42** con *respecto al punto de cruce (cross-over point)*. El umbral de total excepción se establece en **80** *ingresos anuales por innovación como límite inferior*, lo que supone una desviación de **-28.**

**c.** A continuación es necesario traducir estos valores en una métrica de *log odds*. Es una operación sencilla que implica dividir el *logaritmo natural del umbral* de total de pertenencia **(2.944)** por la desviación entre el valor asignado para la *total pertenencia* (**150** *ingresos anuales por innovación*) y el *punto de cruce (cross-over point)* (**108** *ingresos anuales por innovación*): **42.** Esto se traduce en el ratio **2.944 / 42**, o **0.0712**.

**d.** Lo mismo hacemos para la *total exclusión*: se obtiene un radio de **-2.995 / -28**, o **0.1069**. A este radio lo llamaremos *escala*.

**e.** Seguidamente se multiplica la *desviación de cada caso por la escala* asignada. Así se obtiene el producto.

**f.** Finalmente se convierte cada uno de los productos en valores que oscilan entre **0.0 y 1.0** aplicando la siguiente fórmula **(Ragin, 2008, p. 91)**:

**Grado de pertenencia = exp (Producto) / [(1 + exp(Producto) ]**

Donde *exp* representa la *función exponencial* o, en otras palabras, la *función inversa a la función del logaritmo natural*. El resultado final de esta operación genera *grados de pertenencia* en torno a la localización inicial de los *tres valores de referencia* teóricos. Ver la **Tabla 4.7**.

3. *El método indirecto de calibración,* parte de una agrupación previa de los casos que realiza el investigador. A partir de esta clasificación previa de los casos, el investigador recurre a un *modelo logit fraccional* para predecir los valores de regresión de la matriz de datos, que serán usados como indicadores de pertenencia *difusa* (*fuzzy set membership scores*). El software **fsQCA 3.0,** no posibilita la aplicación de este método. Se debe recurrir a **STATA** o a **R**. El método indirecto de calibración requiere que el investigador agrupe los casos de acuerdo con su grado de pertenencia en el conjunto. Para ello se siguen tres pasos:
   - **a.** Primero, el investigador agrupa los casos;
   - **b.** Segundo, se asignan de forma preliminar valores de pertenencia (*membership scores*) a cada grupo; y,
   - **c.** Tercero, se refinan estos valores empleando datos interválicos.

   Con todo esto se consigue que la referencia para fijar los grados de pertenencia sean las agrupaciones que el investigador ha confeccionado previamente y no tanto la localización de los *valores de referencia* teóricos como sucede con el *método directo de calibración*. De esta forma, *la calibración de los casos responde a las interpretaciones cualitativas del investigador* con las que reflejan variaciones entre los datos, en forma de agrupaciones de casos, respaldadas por conocimiento teórico. El software **fsQCA3.0** no permite calibrar con el **método indirecto.** Los investigadores pueden recurrir a **STATA** para ello, el cual permite llevar a cabo el método directo de calibración con el comando **drect.**

Para finalizar, conviene prestar atención a una serie de reglas generales y buenas prácticas para lograr una buena calibración de valores difusos:

1. Primero, *los valores difusos están pensados para capturar variaciones interválicas en términos de conjuntos, por lo que aceptan todo tipo de datos planteados en términos de continuidad en una misma dimensión. Es importante destacar, que un valor de datos difusos no es una*

probabilidad y no responde a principios de análisis estocástico en relación con la prueba de hipótesis probabilísticas. El diseño de investigación continúa siendo *configuracional.*

2. *Tampoco se debe entender un valor de datos difuso de forma multicotómica*. Existe el riesgo de construir un valor *difuso* a partir de varias categorías que pueden no tener cierta continuidad entre ellas. Aunque puede establecerse cierta borrosidad entre, digamos, *empresas verdes y azules*, sería complicado justificar que las *empresas amarillas* representan la *zona intermedia* entre estos dos colores. *No se debe confundir la indefinición entre dos conjuntos con la necesidad de incluir un nuevo conjunto de naturaleza diferente*. En ese caso, **mvQCA** permite asignar valores de tipo *categórico/multicotómico* a este tipo de ejemplos en el que los casos pueden agruparse de acuerdo con si son camisetas verdes, azules, amarillas.
3. *Una matriz de datos requiere ser calibrada*. Valores numéricos no son valores difusos: los valores difusos son binarios y métricos a la vez (Kent, 2009). Se permiten tres tipos de calibración: método de calibración cualitativa, método directo y método indirecto de calibración. Cada uno de estos métodos requiere justificar su uso y la localización de los anclajes teóricos.
4. Se permite la combinación de diferentes estrategias de calibración dentro de una misma matriz de datos. Por ejemplo, se puede encontrar una condición que capture variaciones en el PIB a cuyos valores originales numéricos se les ha aplicado el *método directo de calibración*, otra condición que capture información sobre descentralización de política de innovación por estado, que haya sido calibrada cualitativamente en forma de *cuatro valores difusos,* así como otra condición sobre gasto público de centros de investigación que haya sido calibrada a partir de *tres valores difusos*. Aunque no está pensado para ello, se permite asignar valores dicotómicos si la naturaleza de la condición así lo reclama. Las condiciones con valores *dicotómicos* en matrices de datos en los análisis **fsQCA** acostumbran a ser escasas, pues lo contrario es un sinsentido.

5. Cuando la mayoría de condiciones se presentan en formato *dicotómico* es preferible optar por un análisis **csQCA**. Si la decisión de realizar un análisis **fsQCA** viene determinada por la naturaleza *difusa* del resultado, mientras que el resto de condiciones han sido *dicotomizadas,* entonces es preferible hacer un esfuerzo por *dicotomizar* el resultado y llevar a cabo un análisis **csQCA**.
6. *El resultado tiene que presentar valores de datos difusos* **(Ragin, 2009b, p. 93)**.

## Los conjuntos de datos difusos (*fuzzy-sets*) y las conexiones explícitas

En los apartados anteriores, se presentó la razón por la cual se deben estudiar el conjunto de datos en las ciencias sociales, en especial, las orientadas a la administración de la innovación. Una de las grandes fortalezas del conjunto de *datos difusos*, es que tienen la capacidad de realizar un conjunto de análisis teórico posible, mientras se conservan graduaciones empíricas de conocimiento bien fundamentadas. Esto es, que es posible determinar si un conjunto es un subconjunto de otro (por ejemplo, ¿las empresas innovadoras del clúster calificadas con mayor nivel, son un subconjunto de los países desarrollados?) sin volver a la medición a escala nominal (es decir, conjunto de datos nítidos).

Con un conjunto de datos nítidos, se determina si los casos comparten alguna combinación específica de condiciones comparten los mismos resultados y por lo tanto constituyan un subconjunto de los casos con resultados (una de las dos conexiones explícitas vistas en capítulo 1). El investigador simplemente examina los casos que comparten la combinación relevante de condiciones y evalúa si concuerdan al mostrar el resultado. Esto, se puede ver como una evaluación de la segunda columna de la tabulación cruzada de la presencia / ausencia del resultado en cuestión, vs. la presencia / ausencia de una combinación dada de *condiciones causales*. Ver **Tabla 4.8.**

**Tabla 4.8. Evaluación del conjunto de datos nítidos (*crisp-sets*) de la conexión entre una combinación causal de condiciones y un resultado (la combinación causal es un subconjunto del resultado)**

| | Estatus | Combinación causal | |
|---|---|---|---|
| | | **Ausente** | **Presente** |
| **Resultados** | **Presente** | **C1:** no directamente relevante a la evaluación | **C2:** casos aquí, apoyan el argumento del investigador que esta es una conexión |
| | **Ausente** | **C3:** no directamente relevante a la evaluación | **C4:** debe estar vacío o casi vacío: los casos aquí, debilitan al argumento |

Fuente: Ragin (2008) adaptación propia.

El subconjunto de datos, se indica cuando la celda correspondiente a la presencia de la combinación causal y la ausencia del resultado, está vacía y la celda correspondiente a la presencia de la combinación causal y la presencia del resultado, está poblada con casos. Por supuesto, la **C4** puede no estar completamente vacía. En la investigación orientada a casos, sin embargo, el investigador debe ser capaz de explicar los casos erróneos que pueden encontrar su camino dentro de la celda **C4**.

Obviamente, estos procedimientos no pueden ser duplicados en conjuntos de enlaces difusos. No existe una simple forma de aislar los casos que comparten una combinación específica de *condiciones causales* debido a que cada arreglo de casos de puntuación de pertenencia, pueden ser únicos. Los casos también tienen diferentes grados de pertenencia en los resultados, complicando la evaluación de si convergen en los resultados. Mientras estas propiedades del conjunto de *datos difusos*, hacen difícil duplicar los procedimientos del conjunto de datos nítidos para evaluar al subconjunto de datos, el conjunto de *datos difusos* puede ser evaluado usando álgebra difusa (*fuzzy algebra*). En el conjunto de *datos difusos*, una subconjunto de datos es indicado cuando el puntaje de pertenencia en un subconjunto (por ejemplo, una *condición causal* o combinación de *condiciones causales*) son consistentemente menores que o iguales a sus correspondientes puntajes de pertenencia en otro conjunto (por ejemplo, el de resultados). Para ilustrar lo anterior, considere los datos listados en la **Tabla 4.9.**

**Tabla 4.9. Subconjunto de datos difuso (*fuzzy-sets*) de caso (~U <= W)**

| Empresas innovadoras Clúster | Columnas | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | | 6 | 7 | 8 |
| | Calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado tecnológico de manufactura | Fortaleza resultados consejo de administración | | (~ M) | (A) * (~M) | (A) + ( ~M) |
| | (W) | (A) | (I) | (M) | (U) | ~(U) | | | |
| **A** | **0.6** | 0.8 | 0.6 | 0.4 | 0.6 | **0.4** | 0.6 | 0.6 | 0.8 |
| **B** | **0.6** | 0.6 | 0.2 | 0.2 | 0.8 | **0.2** | 0.8 | 0.6 | 0.8 |
| **C** | **0.2** | 0.6 | 0.4 | 0.2 | 0.8 | **0.2** | 0.8 | 0.6 | 0.8 |
| **D** | **0.8** | 0.6 | 0.8 | 0.2 | 0.2 | **0.8** | 0.8 | 0.6 | 0.8 |
| **E** | **0.6** | 0.6 | 0.8 | 0.4 | 0.4 | **0.6** | 0.6 | 0.6 | 0.6 |
| **F** | **0.8** | 0.2 | 0.6 | 0.8 | 0.8 | **0.2** | 0.2 | 0.2 | 0.2 |
| **G** | **0.6** | 0.4 | 0.8 | 0.2 | 0.6 | **0.4** | 0.8 | 0.4 | 0.8 |
| **H** | **0.8** | 0.6 | 0.4 | 0.2 | 0.4 | **0.6** | 0.8 | 0.6 | 0.8 |
| **I** | **0.2** | 0.6 | 0.4 | 0.6 | 0.8 | **0.2** | 0.4 | 0.4 | 0.6 |
| **J** | **0.0** | 0.8 | 0.4 | 0.8 | 1.0 | **0.0** | 0.2 | 0.2 | 0.8 |
| **K** | **0.4** | 0.6 | 0.6 | 0.8 | 0.6 | **0.4** | 0.2 | 0.2 | 0.6 |
| **L** | **1.0** | 1.0 | 0.8 | 0.4 | 0.2 | **0.8** | 0.6 | 0.6 | 1.0 |

Fuente: Elaboración propia.

La **Tabla 4.9** muestra los puntajes de pertenencia en dos conjuntos de *datos difusos*, el conjunto de empresas innovadoras del clúster con diversos niveles de *calidad de votación por liderazgo tecnológico* (**W**) (solo una = **1**) y el conjunto de empresas innovadoras del clúster con *fortaleza resultados consejo de administración* **(~ U)** no alta, utilizando datos de la **Tabla 4.8**. Observe que los puntajes de pertenencia en la calidad de votos, son consistentemente mayores o igual a los puntajes de fortaleza de resultados consejo de administración no fuertes. Este patrón, es consistente en relación al subconjunto difuso. Si la pertenencia, en la *condición causal* es alta, entonces la pertenencia en el resultado también debe ser alta. Note, sin embargo, que lo contrario no tiene que ser cierto. Es decir, el hecho de que hay casos con pertenencia relativamente baja en la *condición causal* pero sustancial en la pertenencia del resultado (por ejemplo, *empresas innovadoras del clúster* **F** y **B**) no es problemático desde el punto de vista de la teoría de conjuntos porque la expectativa es que puede haber varias formas diferentes de generar una alta pertenencia en el resultado (es decir, hay vías causales para una baja calidad de votación por liderazgo tecnológico, **W,** además del comentado). Casos con puntajes bajos en el *condición causal* (o combinación de condiciones) pero puntajes altos en el resultado, que indica la operación de *condiciones causales* alternativas o alternar combinaciones de *condiciones causales*.

## fsQCA y su relación como técnica basada en conjuntos teóricos

Entender **fsQCA** como una técnica dentro de los métodos de conjuntos teóricos facilita que, salvo algunas excepciones procedimentales, los fundamentos e interpretación de los análisis **fsQCA** sean equiparables a los aplicados a la *lógica booleana*. De hecho, la principal contribución de **Ragin (1987)** es esa: integrar los *conjuntos difusos* en el análisis común de **QCA**, especialmente para lograr un desarrollo paralelo. En este sentido, **Ragin**

**y Pennings (2005, p. 424)** aseguran que los conjuntos difusos mantienen todas las propiedades matemáticas de los conjuntos nítidos (*crisp sets*) y, por lo tanto, permiten trabajar con los diferentes aspectos de casos y examinar relaciones de conjuntos teóricos:

> *"Los conjuntos difusos conservan todas las propiedades matemáticas esenciales de los conjuntos nítidos y, por lo tanto, permiten a los investigadores modelar constelaciones complejas y diversas de aspectos de casos y evaluar las relaciones teóricas de conjuntos [...]. Por ejemplo, los conjuntos difusos permiten la identificación de condiciones necesarias y suficientes por medio del principio de subconjunto: una condición puede interpretarse como* ***necesaria*** *(sus puntuaciones de son consistentemente mayores que el grado de pertenencia en el resultado (la condición es un superconjunto del resultado). Una condición se puede interpretar como suficiente si sus pertenencias son consistentemente menores que el grado de pertenencia en el resultado (la condición es un subconjunto del resultado de interés)".*

Todo esto lleva a que **fsQCA** funcione con los postulados de *diversidad limitada, minimización lógica, soluciones parsimoniosas y complejas, medidas de consistencia y cobertura, contrafácticos, contradicciones, etcétera*. El núcleo duro de lo explicado para **csQCA** y en los capítulos introductorios, tiene utilidad en este capítulo sobre **fsQCA**. Algunas variaciones se detallan en los siguientes párrafos. Los principales aspectos a tratar son las operaciones con conjuntos difusos (negación, **AND** lógico,**OR** lógico), la correspondencia entre espacios de pertenencia y filas de la tabla de verdad, la diversidad limitada, el concepto de subconjuntos y espacios de propiedad, el análisis de condiciones necesarias y el *análisis de condiciones suficientes*.

## Operaciones con conjuntos difusos

La lógica difusa permite conocer el *complemento, o negación,* de un caso en un conjunto de similar forma a la aplicada en **csQCA**. En conjuntos nítidos **(csQCA)** la negación de **(1) es (0)** y viceversa, pues la regla establece que el cálculo de negación del valor de un caso a un conjunto es el resultado de restar ese valor a **(1).** De este modo:

1. Si un caso tiene un valor de **(1)** en **A (A = 1),** esto hace que tenga un valor de **(0)** en **(~ A )(1- 1= 0).** De lo contrario, si tiene un valor de **(0)** en **(A)** significa que tendrá un valor de **(1)** en **(~ A )(1- 0= 1)**. Esta regla se aplica también para *conjuntos difusos*. Un valor de 0.8 en **(Z)** se convierte en **(0.2)** para **(~ Z)** dado el cálculo **(1- 0.8 )= 0.2.**
2. En cuanto a la operación de conjugación **AND** Lógico por la que dos o más conjuntos se combinan, en la lógica difusa se toma el **mínimo** valor de pertenencia de cada caso en los conjuntos combinados. Por ejemplo, si la pertenencia de una empresa al conjunto de *empresas innovadoras en alta tecnología* es de **0.45** y su pertenencia al conjunto de *empresas líderes en precio* es de **0.15**, entonces su pertenencia al conjunto de *empresas innovadoras en alta tecnología* **y** *empresas líderes en precio* adopta el *menor* valor individual de estos dos conjuntos: **0.15.**
3. Para la operación **OR** Lógico, que expresa la unión, también conocida como *disyunción* de dos o más conjuntos, se toma la *máxima* pertenencia del caso a diversos conjuntos que componen la expresión. Si una empresa adquiere una pertenencia *difusa* en el conjunto de *empresas que fabrican productos sofisticados* es de **0.78** y una pertenencia de **0.67** en el conjunto de *empresas con longevidad en el mercado*, su pertenencia en el conjunto *empresas que fabrican productos sofisticados* **o** *empresas que son longevas en el mercado,* asciende a **0.78.**
4. Estos ejemplos permiten conocer las características de esta supuesta empresa. Su pertenencia a empresas *empresas que son líderes en precio* **A** es **de 0.15**, lo que hace que sea *no líder en precio* (valor de **0.85** en **~ A**). Su pertenencia al conjunto *empresas innovadoras de alta tecnología* **(B)** es de **0.45** y su pertenencia al conjunto *empresas no innova-*

*doras de alta tecnología* **(~ B)** es de **0.55**, por lo que es ligeramente una empresa *poco innovadora de alta tecnología.*

5. En el conjunto de *empresas que fabrican productos sofisticados* **C** obtiene un valor de **0.78**, por lo que adquiere un valor de **0.22** en el conjunto **(~ C).** Se trata entonces de una empresa con una *cartera importante de productos sofisticados*. Finalmente, su pertenencia al conjunto **(D),** referido a *empresas con longevidad en el mercado*, es de **0.67**, lo que lleva a una pertenencia de **0.33** en el conjunto **(~ D)**. De ello, se puede concluir que su trayectoria en en el mercado está más allá del promedio.
6. Tomando esto en cuenta, ¿cuál es la *pertenencia* de esta empresa en la configuración **(A * ~C * ~D)** que significa: *empresas líderes en precio* ***y*** *sin productos sofisticados* ***y*** *sin longevidad en el mercado*? La respuesta es **0.15**, pues corresponde al valor mínimo de entre **A (0.15),** ~ **C (0.22) y** ~**D (0.33**).
7. Y ¿en la configuración **(A + ~B + C + ~D)?** que significa: *empresas líderes en precio* ***y sin*** *empresas innovadoras en alta tecnología* ***o*** *empresas que fabrican productos sofisticados* ***y sin*** *empresas longevas en el mercado*? Obviamente será de **0.78** al ser el valor *máximo* de entre las cuatro pertenencias a los conjuntos **A (0.15),** ~ **B (0.55), C (0.78) y** ~ **D (0.33).**

## Análisis de condiciones necesarias

Una condición se convierte en *necesaria, cuando está presente en todos los casos con un mismo resultado, pero su presencia no garantiza dicha ocurrencia.* Esta definición es aplicable a *casos positivos y negativos* por igual **(Ragin, 2000, pp. 205-209)**. *La principal exigencia en conjuntos difusos es que el valor de pertenencia difusa* de *un caso en la condición* **X** *deberá ser igual o superior a su valor de pertenencia difusa en el resultado* **Y (Schneider y Wagemann, 2012, p. 75)**. Esto significa que en la práctica **X** *se trata de un superconjunto de* **Y** *y no uno de sus subconjuntos*, así:

*“Con conjuntos difusos, se señala una posible condición necesaria cuando se puede demostrar que las instancias del resultado en cuestión, constituyen un subconjunto de insistencias de una condición”*. **Ragin, (2009b, p. 110)**

Cabe decir:

**a.** Que el *análisis de necesidad* puede plantearse de forma *individual* para cada una de las condiciones o de forma múltiple, en forma combinatoria **(A * B)** o disyuntiva **(A + B)**. El tamaño desproporcionado de una *disyunción* puede tener efectos indeseados en el *análisis de necesidad*. *Cuanto mayor sea la combinación, por ejemplo,* **(A+B+C+D+F+E),** *más fácil será que los casos, de alguna u otra forma, formen parte de ella*. Se trata de un conjunto claramente más grande que el formado por **(A+B+C) o (D+E).** *Cuanto más grande sea un conjunto, más fácil será que se convierta en un superconjunto del resultado y, en el peor de los casos, una condición trivial*.

**b.** En este caso, la combinación entre dos o más condiciones fuerza a que el valor de pertenencia del caso sea el valor *mínimo* de entre las condiciones combinadas. *En la disyunción se tomará el valor máximo*. Para calcular la *necesidad múltiple*, el software **fsQCA 3.0** permite computar en una nueva variable todas las posibles combinaciones y uniones que el investigador quiera analizar.

**c.** Si bien el criterio tradicional para asignar *estatus de necesidad* en **csQCA** es que el resultado de la consistencia sea de **(1.0)**, en **fsQCA** *difícilmente se puede alcanzar nunca una consistencia de tal magnitud*. En **fsQCA**, el valor de referencia que acostumbra a tomarse gravita en torno a **0.95**, permitiendo interpretaciones menos conservadoras a partir de **0.85**. No obstante, lo cierto es que la comunidad científica tiende a imponer un criterio férreo al respecto y, en el caso de querer el investigador rebajar el *valor de referencia por debajo de* **0.95**, deben ser aportados *análisis complementarios*, tales como el análisis visual con **XY Plots** o el valor **PRI**, que justifiquen cualitativamente su decisión.

De entre estas herramientas, los **XY Plots** cumplen una *cuádruple función* en el momento de calcular condiciones necesarias:

1. Primero, obtener un primer *análisis visual* de la distribución de casos y sus correspondientes valores *difusos*. De esta forma se comprueba visualmente el grado de dispersión y concentración entre los casos para, por ejemplo, revisar el modelo de calibración.
2. Segundo, *comprobar si la condición* **X** *cumple los criterios de necesidad para con* **Y**. *Para ello todos* los casos *deberán estar sobre o por debajo de la diagonal (o dentro del triángulo inferior)*. Esto es ejemplo de que, asumiendo que los casos pertenecen a ambos conjuntos, la pertenencia al conjunto **X** es mayor o igual a su pertenencia en el conjunto Y, lo que gráficamente conduce a que la frontera de **X** (*triángulo inferior*) sobrepase la frontera establecida por **Y** (*triángulo superior*). Si no ocurre esto, el argumento a favor de la consideración de necesidad quedará debilitado: asumiendo pertenencia a ambos conjuntos, *la pertenencia* a **X** *sería infelior a la pertenencia a* **Y** *haciendo que* **X** *se convirtiera en un subconjunto de* **Y.**
3. Tercero, *analizar la trivialidad de la condición necesaria*. Una condición es trivial cuando se convierte en un *superconjunto* de cualquier resultado. Decir que para llegar a ser directivo CEO, **D** es necesario haber defendido una tesis doctoral **T** es una afirmación tan trivial como considerar que se trata de personas que han completado estudios, de licenciatura o de grado, superiores **E.** *Escenarios de trivialidad* ocurren cuando muchos casos se agrupan *cerca del eje vertical derecho* del **XY Plot (Schneider y Wagemann, 2012, p. 146)**. En realidad, el *análisis de condiciones necesaria*s busca *condiciones relevantes que no triviales* **(Ragin, 2008)**. *Las condiciones necesarias relevantes se conocen como condiciones no triviales por oposición a las condiciones triviales.* Una manera de inspeccionar *la trivialidad de las condiciones necesarias* es prestar atención a la medida de *cobertura* ofrecida durante el *análisis de condiciones necesarias*. Esta medida expresa *la proporción de casos que están presentes en la condición* **X** *y en el resultado entre los casos*

*que presentan la condición* **Y**. Dicho de otra forma, la *cobertura mide cuán pequeño es el conjunto de* **Y**. *En relación con el conjunto de* **X**. Si los conjuntos **X** e **Y** son parecidos en tamaño, entonces la cobertura de **(X)** como una condición necesaria de **(Y)** es alta **(Schneider y Wagemann, 2012, p. 144)**. **Bol y Luppi (2013)** plantean que *una cobertura de necesidad extremadamente baja indica que la condición es constante para el universo de casos* y, por ello, irrelevante para explicar el resultado.

**4.** Y, cuarto, conocer si la *calibración de la matriz de datos* padece un problema de medición imprecisa, por lo que sería conveniente redefinir los *criterios de calibración* **(Ragin, 2000, p. 223)**.

Finalmente, como sucede para **csQCA**, en **fsQCA** el *análisis de condiciones necesarias* corre el riesgo de encontrar *falsas condiciones necesarias* y de *esconder verdaderas condiciones necesarias* a lo largo del análisis de *suficiencia.* Las particularidades *difusas* de este análisis son explicadas en las siguientes líneas:

**a.** En lo que respecta a las *condiciones necesarias*, los riesgos que se apuntan son causados, como sucede para **csQCA**, por el tratamiento que se haga de los *residuales* y por la posibilidad de encontrar *filas inconsistentes* en la *tabla de verdad*.

**b.** En cuanto a los *primeros,* si una *condición necesaria* desaparece por haber planteado conjeturas incoherentes en el momento de definir expectativas direccionales (para obtener la *solución intermedia*), la solución es fácil: *no se deben hacer conjeturas que contradigan la afirmación de que una condición es necesaria* **(Schneider y Wagemann, 2012, pp. 223-224)**, es decir*, los residuales que no contengan la condición necesaria no deben utilizarse para la minimización del resultado.*

**c.** En cuanto a las segundas, que tanto pueden generar falsas condiciones necesarias como ocultar condiciones que superarían el test de necesidad, *se debe evitar seleccionar umbrales de consistencia demasiado bondadosos* en el momento de definir los *casos positivos y los casos negativos,*

que permitan la inclusión *de configuracione*s que durante la *minimización* hagan desaparecer condiciones verdaderamente necesarias.

## Análisis de condiciones suficientes

Como se ha mencionado en otros capítulos:

**a.** El *análisis de condiciones suficientes* se fundamenta en la idea de que *si una o varias condiciones combinadas exhiben el mismo resultado, entonces estas condiciones forman subconjuntos dentro del (super) conjunto* **Y**.

**b.** La idea de *causalidad múltiple* establece que *más de un subconjunto puede tener cabida dentro del conjunto formado por* **Y**, *por lo que varias combinaciones explican la ocurrencia de* **Y**.

**c.** También, se ha comentado que estos *subconjuntos* pueden llegar a *solaparse*, lo que permitiría detectar configuraciones capaces de *explicar la causalidad de forma aún más parsimoniosa*.

**d.** En caso de *sobredeterminación*, cada conjunto explicaría unos casos de forma exclusiva *(cobertura única)*, mientras que el resto formarían parte de varios subconjuntos *(cobertura bruta)* a la vez.

**e.** El tamaño conjunto de todos los *subconjuntos* por ejemplo **(A+ ~ B * C + G * ~ E → Y)** se encuentra en el **solution term**. Esto, que tiene aplicación en **csQCA**, también la tiene para **fsQCA** con una ligera diferencia: *en conjuntos difusos los investigadores utilizan la relación aritmética: menos que o igual a*, para definir las relaciones de suficiencia **(Ragin, 2000, p. 235)**. La razón para ello no es otra que la recurrente apelación a la *pertenencia difusa* de los casos a los conjuntos establecidos.

**f.** Una condición podrá obtener *suficiencia causal* cuando se demuestre que es un *subconjunto del resultado*, es decir, que una condición con un valor de **0.64** será suficiente para explicar un resultado que sea igual o mayor que esta **(>=0.64).** De lo contrario, **(X)** sería un *superconjunto*

de **Y** dado que **Y** estaría dentro de **X**. Si **X** está presente siempre que **Y** ocurre, **X** se vuelve una condición necesaria.

**g.** Igual que en **csQCA**, lo interesante en el análisis de condiciones suficientes es conocer la *consistencia* (relación *entre el número de casos de* **X** *dentro y fuera del conjunto* **Y**, o *relación entre el número de casos dentro del triángulo superior y el número de casos fuera de él*) y la *cobertura* (tamaño del subconjunto **X** dentro del conjunto **Y**, o cuántos casos con ocurrencia de **Y** es capaz de explicar **X**).

**h.** Gráficamente, a partir de los **XY Plots**, se detecta que *una condición cumple con el criterio de suficiencia cuando todos los casos caen dentro del triángulo superior* **(XY).** Al estar todos los casos *por encima de o sobre la diagonal se entiende que todos ellos se encuentran dentro de* **Y.**

**i.** Uno o varios casos por *debajo de la diagonal* indican que el tamaño de la **X** sobrepasa los límites de **Y**, evidenciando la existencia de casos (**X, ~ Y)**. Recuérdese que *el eje lateral izquierdo ubica el resultado mientras que el horizontal inferior indica el espacio que ocupa la condición* **X** *y la diagonal es la frontera entre ellos.*

**j.** En la práctica, el análisis de condiciones suficientes en **fsQCA** sigue las operaciones de *negación y conjunción* para identificar las posibles expresiones causales.

**k.** La **Tabla 4.10** muestra una *matriz de datos hipotética* a partir de dos condiciones **(A y B)** y un resultado **R.**

**Tabla 4.10. Matriz hipotética de datos.**
**Prueba de suficiencia con valores difusos**

| Casos | Columna | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (A) | (B) | (~ A) | (~ B) | (A * B ) | (~ A * B) | (A * ~ B) | (~ A * ~ B) | R |
| 1 | 0.08 | 0.27 | 0.92 | 0.73 | 0.08 | 0.27 | 0.08 | 0.73 | 0.20 |
| 2 | 0.17 | 0.35 | 0.83 | 0.65 | 0.17 | 0.35 | 0.17 | 0.65 | 0.28 |
| 3 | 0.89 | 0.46 | 0.11 | 0.54 | 0.46 | 0.11 | 0.54 | 0.11 | 0.51 |
| 4 | 0.67 | 0.70 | 0.33 | 0.30 | 0.67 | 0.33 | 0.30 | 0.30 | 0.82 |
| 5 | 0.53 | 0.91 | 0,47 | 0.09 | 0.53 | 0.47 | 0.09 | 0.09 | 0.53 |
| 6 | 0.47 | 0.33 | 0.53 | 0.67 | 0.33 | 0.33 | 0.47 | 0.53 | 0.41 |
| 7 | 0.65 | 0.75 | 0.35 | 0.25 | 0.65 | 0.35 | 0.25 | 0.25 | 0.69 |
| 8 | 0.70 | 0.30 | 0.30 | 0.70 | 0.30 | 0.30 | 0.70 | 0.30 | 0.31 |

Fuente: Elaboración propia.

**l.** Para este ejemplo, cada una de las filas corresponde a un caso. Se observa que a partir de las condiciones **(A y B)** se obtienen otras expresiones como **(~ A y ~ B)**, cuyos valores resultan de restar el valor de **(A y B)** a 1 respectivamente. Otras expresiones lógicas son resultado de las posibles combinaciones de **(A y B),** a saber, **(A*B), (~ A * B), (A * ~ B)** y **(~ A * ~ B).** Los valores que toman estas expresiones corresponden al *valor mínimo* de sus componentes. Las *celdas grisáceas indican cuándo el valor de la condición cumple con el criterio de suficiencia*: *ser igual a o menor que el valor del resultado.*

**m.** Solamente la expresión **(A * B)** cumple con este criterio en su totalidad, mientras que el resto de expresiones muestran valores superiores en algún momento. Resultado de esta simple prueba se puede asegurar que **(A * B)** *cumple el criterio de suficiencia* para el resultado, siendo un subconjunto de él **(AB → Y).**

**n.** Sin embargo, **Schneider y Wagemann (2012, p. 69)** valoran *que encontrar relaciones perfectas de suficiencia en conjuntos difusos es manifiestamente más complicado que en* **csQCA**, por lo que se requiere cierto grado de flexibilidad en el establecimiento de los *umbrales de frecuencia (o consistencia)*. En otras palabras, existe cierta flexibilidad

para permitir que algunos casos tuvieran valores en la **X** mayores que en la **Y** y, aun así, considerar que **X** cumple con los criterios de suficiencia. Esta flexibilidad adquiere relevancia en el momento de especificar los umbrales de frecuencia en las *tablas de verda*d. Es en ese momento cuando el investigador se ve obligado a especificar el estatus *positivo o negativo* de un caso, codificando las combinaciones con un **(1)** o un **(0)**, a la vez que se perfilan los *residuales y las contradicciones*. En efecto, se trata de diferenciar entre combinaciones causales consistentes de las que no lo son.

**o.** En **csQCA**, cuando una expresión cuenta con evidencia empírica y no alberga *contradicciones*, el valor de la consistencia alcanza el **(1)**. Las *contradicciones* se descartan o se intentan resolver, pero en todo caso su consistencia será *sensiblemente inferior a* **(1)** *(para dos casos, uno con resultado positivo y otro con resultado negativo, la consistencia es* **0.5***).*

Para **fsQCA**, el investigador debe analizar cuidadosamente, al menos, tres cuestiones:

1. Primero, identificar las configuraciones causales consistentes, esto es, configuraciones que cuentan con casos que tengan valores de pertenencia mayores a **0.5** (*cross-over point*), significando con ello que estos casos pertenecientes a **(X)**, están a su vez *más dentro que fuera del resultado*. La primera respuesta sería codificar como *configuraciones inconsistentes,* con un código **(0)**, las configuraciones con un valor menor que **(0.5)**. Se pueden encontrar casos que cumplan el resultado pero que formen parte de configuraciones causales con una *baja consistencia*. **Ragin (2009b, p. 109)** apunta que, en estos casos, estas configuraciones se asemejan mucho a las *configuraciones contradictorias* en **csQCA**, pudiéndose aplicar soluciones similares para resolver su *contradicción* y *aumentar el nivel de consistencia.*
2. Segundo, **Ragin (2009b, pp. 106-107)** plantea que *es necesario disponer de un número mínimo de casos dentro de las configuraciones causales consistentes para distinguir entre configuraciones relevantes e*

*irrelevantes*. No cabe duda de que una configuración con muchos casos por encima de **0.5** claramente determina una *relación de subconjunto* con respecto al resultado. *Si son pocos los casos dentro de una configuración causal consistente, entonces aparecen dudas acerca de su relevancia.* Determinar el significado de muchos casos depende del universo de casos. Si la matriz de datos se compone de una *veintena de casos,* obviamente la variedad de casos en, por ejemplo, diez configuraciones consistentes será *moderada*. Así:

**a.** *Para matrices con pocos casos, el umbral de casos para cada configuración puede establecerse en* **1 o 2**.

**b.** *Para matrices con centenares de casos, este umbral deberá ser sustancialmente mayor* **Ragin (2009b, p. 107)**, apunta:

> *"...cuando el número total de casos en un estudio es grande, el problema no es qué combinaciones tienen instancias (por ejemplo, al menos un caso con mayor de 0.5 miembros), sino qué combinaciones tienen suficientes instancias para esperar a realizar una evaluación de su posible relación de subconjunto con el resultado…"*

**c.** Cuando no hay evidencia empírica en las configuraciones, el paso a seguir es considerarlas como *residuales.*

**3.** Tercero, en referencia a la identificación de las *configuraciones causales consistentes relevantes*, es necesario tener en cuenta el efecto que tiene establecer como referencia un *umbral de consistencia de* **>0.5.** Estableciendo este umbral, *casi la mitad de la evidencia empírica contradice el criterio de suficiencia*. Incluso si se eleva ese umbral hasta **>0.75** *se estaría aceptando que una cuarta parte de los casos pueden estar fuera del subconjunto*. **Ragin (2008, p. 136)** propone que el valor orientativo para indicar la *consistencia sustancial debe establecerse en torno a* **0.85**, haciendo que las configuraciones con valores inferiores se caractericen por una cierta *inconsistencia sustancial*. En la actualidad incluso se observa una *tendencia a fijar umbrales más exigentes superiores a* **0.9**. No obstante, la imposición de un valor de referencia prefijado para

todo tipo de investigaciones es una *mala práctica* **(Schneider y Wagemann, 2012, p. 127)**. *A menudo la imposición del umbral es resultado de la existencia de un hueco destacado entre niveles de consistencia* (por ejemplo, cuando se encuentra una variación entre **0.84** y **0.72** o similar). De otra parte, la justificación del umbral de consistencia puede tener un respaldo teórico.

En relación con el parámetro de *cobertura*, más allá de conocer cuál es el peso individual y combinado de un subconjunto cualquiera en el total de casos dentro de **Y**, *lo importante es saber detectar algunas de las razones que hacen incrementar o disminuir la cobertura de las condiciones suficientes.* Pero antes que nada, no es menos importante saber que solamente tiene sentido calcular la *cobertura* de las condiciones que anteriormente hayan sido definidas como *consistentemente suficientes*. *No hay motivo aparente para calcular la cobertura de una condición que no es un subconjunto de* **Y**, por lo que:

1. La primera lección a tener en cuenta es que *un valor de cobertura pequeño no es sinónimo de estar incurriendo en algún tipo de error procedimental o metodológico*; simplemente se está indicando que la configuración cubre una pequeña porción del resultado. Lo importante será averiguar el estatus de subconjunto de esta configuración con respecto al resultado, es decir, si **X** se encuentra mayoritaria y consistentemente dentro de **Y**.
2. Puede haber cierta preocupación cuando el análisis incorpora un *número excesivo de condicione*s, pues posibilita encontrar expresiones lógicas del tipo **(A * B * C * D * E * F).** La pertenencia a conjuntos compuestos por muchas condiciones se vuelve más difícil y, por ende, se trata de *conjuntos pequeños con potencialmente escasa cobertura* que fácilmente pueden ser subconjuntos de **Y** (es más fácil que cumplan con el criterio de consistencia). De lo contrario, *cuanto más espacio de* **Y** *sea capaz de cubrir* **X**, *mayor cobertura*.

3. En términos gráficos es fácil reconocer que la concentración de valores de **X** en un área determinada dentro del conjunto **Y** (por ejemplo, en el lateral izquierdo) se traduce en que el tamaño de **X** no es equiparable al de **Y**. *Una dispersión equilibrada a lo largo del triángulo superior beneficia al valor de cobertura*. Este hecho está estrechamente relacionado con el siguiente argumento: *la cobertura de una condición suficiente se incrementa cuanto más cerca esté de la diagonal. La proximidad a la diagonal es relevante por una sencilla razón: cuando un caso está cerca de la diagonal, su pertenencia difusa en* **X** *es similar a su pertenencia difusa en* **Y**. *Los casos alejados de la diagonal presentan una pertenencia en* **Y** *mayor que en* **X:** *son buenos ejemplos de* **Y** *pero son ejemplos débiles de* **X**. Sin embargo, obtener soluciones con un solo patrón causal en el *análisis de condiciones suficientes* es una posibilidad, que no la única.
4. Soluciones equifinales, resultado del proceso de minimización lógica, que en **fsQCA**, como en **csQCA**, se realiza mediante el algoritmo **Quine-McCluskey**, obligan a preguntarse *cuánto suman las diferentes coberturas de cada una de las rutas que explican el resultado*. Cada una de estas rutas ocupa un espacio dentro de **Y**, a veces bien definido y exclusivo, a veces solapado con otras rutas. Estos espacios pueden explicar una pequeña porción del resultado, lo que genera una *cobertura global relativamente baja.* De lo contrario, *una cobertura global alta se produce cuando las diferentes rutas son capaces de cubrir una alta cantidad del tamaño del resultado*. Esto no implica que todas las rutas contribuyan de igual manera a cubrir espacios de **Y**: la cobertura global es la suma de coberturas grandes y pequeñas, únicas y solapadas.
5. Por norma general, estos planteamientos provienen de la lógica **QCA** y, por lo tanto, no son de invención *ad hoc* para los conjuntos difusos. El mismo tipo de soluciones complejas, intermedias y parsimoniosas se pueden obtener en **fsQCA** como en **csQCA** bajo los mismos supuestos de *inclusión* y *exclusión de residuales*, así como la imposición de expectativas direccionales. El sentido y el significado de *las implicaciones principales* son también compartidos con **csQCA**: *el objetivo es encon-*

*trar la solución más parsimoniosa posible*. A lo sumo, se encuentra una ligera diferencia cuando se muestran los casos cubiertos por cada ruta. El software **fsQCA 3.0** indica siempre los casos con una *pertenencia superior* a **0.5** en dicha ruta acompañados de dos dígitos entrecomillados del tipo **(0.95, 1).** Esta información se refiere a la pertenencia del caso a **(X, Y).**

6. Al ser mayoritariamente configuraciones combinadas del tipo **(A * B * ~ C)**, el valor del caso en **X** será el menor de entre las condiciones que forman la ruta, siempre superior a **0.5**. Para el valor en **Y** es lógico pensar que serán valores muy cercanos a **(1)**, pues es resultado de la *imposición de altos umbrales de consistencia* para las expresiones que cumplen con el *criterio de suficiencia*.
7. Por estas semejanzas en el procedimiento y diferencias en la obtención de valores, se vuelve más oportuno, en el marco de un entorno metodológico, centrarse en cómo se obtienen las medidas de cobertura únicas y brutas en el análisis de conjuntos difusos. Se puede consultar las reglas de minimización y demás critetios en el capítulo dedicado a **csQCA** dado que, como decimos, son idénticos.
8. Tomando los datos de la tabla anterior, y suponiendo ahora que las configuraciones **(A * B)** y **(~ A)** son subconjuntos hipotéticos consistentes con el resultado, se procede a calcular la *cobertura conjunta* (**solution coverage**) y la *consistencia conjunta* (**solution consistency**). Lo *primero* expresa el *tamaño global de los varios subconjuntos dentro de* **Y**, y lo *segundo* se refiere a la *pertenencia de estos subconjuntos dentro de* **Y**, por lo que:
   - **a.** El primer paso es, para cada caso, extraer el valor mayor entre **(A * B + ~ A)**, que serán sumados **(a).** La tercera columna de la **Tabla 4.11** recoge los valores máximos que suman un valor de **4.89**.

**Tabla 4.11. Matriz hipotética de datos.**
**Cálculo de la cobertura conjunta y la consistencia conjunta en fsQCA**

| Casos | Columnas | | | | |
|---|---|---|---|---|---|
| | (A * B) | (~ A) | (A * B) + (~ A) | (R) | min (A * B + ~ A * B, R ) |
| **1** | 0.08 | 0.92 | 0.92 | 0.20 | 0.20 |
| **2** | 0.17 | 0.83 | 0.83 | 0.28 | 0.28 |
| **3** | 0.46 | 0.11 | 0.46 | 0.51 | 0.46 |
| **4** | 0.67 | 0.33 | 0.67 | 0.82 | 0.67 |
| **5** | 0.53 | 0,47 | 0.53 | 0.53 | 0.53 |
| **6** | 0.33 | 0.53 | 0.53 | 0.41 | 0.41 |
| **7** | 0.65 | 0.35 | 0.65 | 0.69 | 0.65 |
| **8** | 0.30 | 0.30 | 0.30 | 0.31 | 0.30 |
| **Suma** | — | — | **4.89 (a)** | **3.75(b)** | **3.50 (c)** |

Fuente: Elaboración propia.

**b.** El segundo paso es sumar todos los valores individuales del resultado **(b).** El resultado es **3.75**.

**c.** Por último se anotan, para cada caso, los valores mínimos de **(AB + ~ A)**, y **R**, que serán sumados **(c)**.

**d.** La *cobertura conjunta* se obtiene al dividir **(c)** / **(b),** lo que resulta en una cobertura de **3.50 / 3.75 = 0.9333**.

**e.** La *consistencia conjunta* es el resultado de dividir **(c) / (a),** que es igual a **0.715.**

**9.** Para calcular la *cobertura única* y la *cobertura bruta* de cada una de las soluciones, el procedimiento es parecido al anterior. Ahora interesa centrarse en la relación individual de cada condición con el resultado, así como en posibles solapamientos, por lo que:

**a.** El primer paso será obtener la suma de los valores de cada una de las dos condiciones suficientes por separado. Las *columnas 1 y 2* de la **Tabla 4.12** reflejan el resultado de la suma.

**Tabla 4.12. Matriz hipotética de datos.**
**Cálculo de la cobertura única y la cobertura bruta en fsQCA**

| Casos | Columnas | | | | |
|---|---|---|---|---|---|
| | (A * B) | (~ A) | (R) | min (A * B, R ) | min (~ A, R ) |
| **1** | 0.08 | 0.92 | 0.20 | 0.08 | 0.20 |
| **2** | 0.17 | 0.83 | 0.28 | 0.17 | 0.28 |
| **3** | 0.46 | 0.11 | 0.51 | 0.46 | 0.11 |
| **4** | 0.67 | 0.33 | 0.82 | 0.67 | 0.33 |
| **5** | 0.53 | 0,47 | 0.53 | 0.53 | 0.47 |
| **6** | 0.33 | 0.53 | 0.41 | 0.33 | 0.41 |
| **7** | 0.65 | 0.35 | 0.69 | 0.65 | 0.35 |
| **8** | 0.30 | 0.30 | 0.31 | 0.30 | 0.30 |
| Suma | 3.19 (a) | 3.84 (b) | 3.75 (c) | 3.19(d) | 2.45 (e) |

Fuente: Elaboración propia.

Además es necesaria la suma de los valores del resultado, mostrado en la *columna 3*. La *columna 4 y 5* muestran la suma de los valores mínimos de cada una de las condiciones y el resultado.

**b.** La *consistencia* de **(A * B)** en **R** se obtiene al dividir **(d)** por **(a)** que genera un resultado de **3.19 / 3.19 = l.**

**c.** La *cobertura bruta* de **(A * B)** se obtiene dividiendo **(d) / (c).** El resultado es **3.19 / 3.75 = 0.85.**

**d.** Para obtener la *cobertura única* de **(A * B)** se necesita conocer la *cobertura conjunta* del modelo. Esta ha sido anteriormente calculada y es igual a a **0.9333**.

**e.** La *cobertura única* de **(A * B)** se obtiene al restar la *cobertura bruta del resto de condiciones* **(~ A)** a la cobertura conjunta: **0.9333 - 0.65 = 0.283**.

**10.** Para obtener los parámetros de consistencia y cobertura para **(~ A)** basta con seleccionar otros sumandos, así:

**a.** La consistencia de **(~ A)** en **R** resulta de la división **(e) / (b),** esto es, **2.45 / 3.84 = 0.64**.

**b.** La cobertura bruta de **(~ A)** resulta de dividir **(e) / (c)** de la siguiente manera **2.45 / 3.75 = 0.65**.

**c.** Se obtiene la cobertura única de **(~ A)** al restar la *cobertura bruta* de **(A * B)** a la *cobertura conjunta*: **0.9333 - 0.85 = 0.083.**

**11.** Todos estos cálculos permiten obtener información precisa sobre el análisis de necesidades suficientes de este ejemplo hipotético:

- Condición **(A * B):** *cobertura única* **(0.283)**, *cobertura* bruta **(0.85),** *consistencia* **(1).**
- Condición **(~ A)**: *cobertura única* **(0.083),** *cobertura bruta* **(0.65),** consistencia **(0.64).**
- *Cobertura conjunta:* **(0.9333).**
- *Consistencia conjunta***: (0.715).**

Nótese que los valores de *cobertura única y bruta* obtenidos en ambas soluciones *no son coincidentes*. La razón es bien sencilla: existe sobre-determinación. Cuando los valores de la cobertura única y bruta coinciden es sinónimo de que la solución no está solapada y los casos que explica no se encuentran en otros subconjuntos del resultado.

Un aspecto especialmente útil del pensamiento *configuracional*, es su atención. a la *complejidad causal*, la cual se define como una situación en la cual, un resultado dado, puede provenir de varias combinaciones diferentes de *condiciones causales*, de diferentes trayectorias causales. Por ejemplo, un investigador puede tener buenas razones para sospechar, que varias trayectorias causales diferentes, conducen a un posicionamiento dentro del reporte de índice de innovación global **(GII, 2019)**, así:

**1.** Al examinar los casos con diferentes combinaciones de condiciones de causalidad relevantes, es posible identificar las trayectorias causales decisivas y de este modo desentrañar la *complejidad causal*.

**2.** Como se explica en el capítulo 1 y basados en **Ragin (1987, 2000)**, la herramienta clave para el análisis sistemático de la *complejidad causal*, es la tabla de verdad. Las *tablas de verdad* nítidas (*crisp truth tables*), enumeran las combinaciones lógicamente posibles de las *condiciones*

*causales* dicotómicas, (por ejemplo, ver **Tabla 7.3**, con presencia / ausencia de los factores: *instituciones* **N,** y/o c*apital humano e inves-tigación* **H,** y/o infraestructura **I,** y/o *sofisticación del mercado* **M**, y/o s*ofisticación de negocio* **B,** y/o c*onocimiento y tecnologías producidas* **K**, y/o *creatividad producida* **C**, junto con el resultado exhibido por los casos que se ajustan a cada combinación de *condiciones causales* (por ejemplo, si el nivel de innovación más destacado por su mayor puntua-ción. está constantemente presente entre los casos que muestran para cada combinación de condiciones).

**3.** Una *tabla de verdad,* usando la *trayectoria causal* de cuatro condiciones tendría dieciséis filas, una para cada combinación lógicamente posible de *condiciones causales*. En *tablas de verdad* más complejas, las filas (cada una representa una combinación diferente de *condiciones causales*) pueden ser numerosas, debido a que el número de combinaciones causales es una función exponencial del número de *condiciones causales* (número de combinaciones = $\mathbf{2}^{k}$, donde ***k*** es el número de *condiciones causales*). En efecto, una *tabla de verdad* de datos nítidos convierte a ***k*** en presencia / ausencia de *condiciones causales* dentro de $\mathbf{2}^{k}$ *con iguraciones* como se ha ilustrado en varias tablas de este trabajo, como la **Tabla 4.12a.**

**Tabla 4.12a. Tabla de verdad que relaciona las variables de la administración del conocimiento del consumidor CKMS, con los procesos de innovación, INNOV**

| Combinación | INNOV | CKMADI | CKMS | CKMOSK | CKMSEP |
|---|---|---|---|---|---|
| 1 | — | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 1 |
| 3 | 0 | 0 | 0 | 1 | 0 |
| 4 | — | 0 | 0 | 1 | 1 |
| 5 | 0 | 0 | 1 | 0 | 0 |
| 6 | 0 | 0 | 1 | 0 | 1 |
| 7 | — | 0 | 1 | 1 | 0 |
| 8 | 1 | 0 | 1 | 1 | 1 |
| 9 | 0 | 1 | 0 | 0 | 0 |
| 10 | — | 1 | 0 | 0 | 1 |
| 11 | — | 1 | 0 | 1 | 0 |
| 12 | — | 1 | 0 | 1 | 1 |
| 13 | 1 | 1 | 1 | 0 | 0 |
| 14 | 1 | 1 | 1 | 0 | 1 |
| 15 | — | 1 | 1 | 1 | 0 |
| 16 | 1 | 1 | 1 | 1 | 1 |

Notas:

0. Los resultados marcados con "–" indican que no existe una comprobación empírica con las combinaciones listadas en el renglón.
1. Es Sí
2. Es No

Fuente: Mejía-Trejo (et al., 2016).

**4.** El enfoque de la *tabla de verdad* sobre la *complejidad causal* es mucho más riguroso y exigente, donde la preocupación en se fijó en una trayectoria causal dada y las diversas versiones simplificadas de esa trayectoria causal, que podrían construirse a partir de subconjuntos de condiciones especificadas en la trayectoria causal inicial. Por el contrario, el enfoque de la *tabla de verdad*, se basa en todas las combinaciones lógicamente posibles de las *condiciones causales* especificadas por el investigador,

quien generalmente incluye condiciones extraídas de diferentes perspectivas y, por lo tanto, diferentes trayectorias causales.

5. Además, en lugar de analizar diferentes subconjuntos de un conjunto dado de condiciones (con condiciones presentes o *irrelevantes*), el enfoque de la *tabla de verdad*, considera todas las combinaciones de *condiciones causales* lógicamente posibles, tomando en cuenta, tanto su presencia como su ausencia.
6. Por lo tanto, el enfoque de la *tabla de verdad*, permite la posibilidad de que diferentes trayectorias causales puedan operar cuando una condición dada (por ejemplo, la percepción de sofisticacion de mercado **M** en el ejemplo de posicionamiento de un país en el *Global Innovation Index***(GII , 2019)** está presente vs. ausente (*en oposición a irrelevante*). Por ejemplo, una trayectoria causal para el posicionamiento en el índice de innovación global, podría requerir la ausencia de un factor como la *sofisticacion de mercado* **M** en un país emergente, para explicar ya sea su posicionamiento o cómo mejorarlo.
7. Examinando todas las combinaciones lógicamente posibles de las *condiciones causales*, hace posible proponer y construir experimentos con contraste de diseño (donde solo hay una *condición causal* a un tiempo que puede variar) y, por lo tanto, ofrece un análisis exhaustivo de los efectos de *condiciones causales relevantes*. En efecto, el impacto de cada causa se examina en todos los contextos lógicamente posibles (las *configuraciones* $\mathbf{2}^{k}$ de condiciones).
8. El objetivo de la construcción de la *tabla de verdad*, es identificar conexiones explícitas entre combinaciones de *condiciones causales* y sus resultados. Utilizando la *tabla de verdad*, es posible evaluar el ser suficiente de todas las combinaciones lógicamente posibles, de condiciones de presencia / ausencia ($\mathbf{2}^{k}$ *configuraciones* causales) que se pueden construir, a partir de un conjunto dado de *condiciones causales* ***k***. Las combinaciones que pasan el ser suficiente, son entonces lógicamente simplificadas de abajo hacia arriba.

Por ejemplo, en el análisis de condiciones vinculado al posicionamiento del *Global Innovation Index* **(GII, 2019)**, si los casos con las cuatro condiciones presentan s*ofisticación de negocios* **B** y los casos con tres de las cuatro condiciones presentes (*y una ausente*) también experimentan *sofisticación de negocios* **B**, entonces:

**1.** El investigador puede concluir que la *condición causal* que varía a través de estas dos combinaciones, es *irrelevante* en este contexto. **(Ragin 1987, 2000)**. Las *tablas de verdad*, también disciplinan el proceso de aprendizaje sobre los casos y el esfuerzo de generalizar sobre ellos. Por ejemplo, suponga una *tabla de verdad* para un posicionmiento del *Global Innovation Index* **(GII, 2019)** para el caso de México, basada en las conclusiones extraídas de los capítulos anteriores, revela *inconsistencias* sustanciales en la fila de México, es decir, supongamos que hay varios casos de países en la fila, que no se exhibieron como México hizo.
**2.** Esta *inconsistencia* de los resultados, son evidencias para el investigador, de que son necesarios más estudios de casos en profundidad. Por ejemplo, al comparar los casos en este fila sin la *sofisticación de negocios* **B** con los que sí la exhiban, haría posible elaborar la trayectoria causal. Supongamos, que esta comparación revela que los casos que carecen de *sofisticación de negocios* **B,** la gran mayoría presentaban falta de tecnología de soporte para mercadotecnia digital y negocios electrónicos.
**3.** Este componente (*la ausencia de tecnologías de soporte*) podría entonces agregarse a la trayectoria causal, y la *tabla de verdad* podría reformularse en consecuencia con cinco *condiciones causales*. Note que hubiera sido difícil de saber, basado en el conocimiento solo del caso de México, que este factor (*la ausencia de tecnologías de soporte*), fuera importante en la trayectoria causal porque está (hipotéticamente) ausente en México y en casos similares al de México. Este punto subraya el valor del análisis comparativo, de manera más general, ya que a menudo es difícil identificar componentes causales que deben estar ausentes cuando se estudian solo instancias positivas de un resultado.

## Conjunto de datos difusos (*fuzzy-sets*) y tablas de verdad

Este capítulo pretende relacionar el conjunto de *datos difusos*, las *configuraciones* causales y el análisis de la *tabla de verdad*, demostrando cómo construir una *tabla de verdad* Booleana convencional a partir de un conjunto de datos difuso y posteriormente, cómo usarla para desentrañar la *complejidad causal*. Esta técnica:

1. Aprovecha al máximo, las gradaciones en la pertenencia del conjunto central, para la constitución de conjuntos de *datos difusos* y no se basa precisamente en una *dicotomización* de dichas puntuaciones de pertenencia del conjunto difuso. Para ilustrar estos procedimientos, el capítulo utiliza de nueva cuenta, los datos hipotéticos introducidos en el **Tabla 4.5** y que se muestran en la **Tabla 4.13.**

**Tabla 4.13. Conjunto de datos difusos (*fuzzy-sets*) para el caso de decisión de clúster de innovación**

| Empresas innovadoras Clúster | Columnas | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | Calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado tecnológico de manufactura | Fortaleza resultados consejo de administración |
| | W | A | I | M | U |
| A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 |
| C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 |
| D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 |
| E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 |
| F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 |
| G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 |
| H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 |
| I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 |
| J | 0.0 | 0.8 | 0.4 | 0.8 | 1.0 |
| K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 |
| L | 1.0 | 1.0 | 0.8 | 0.4 | 0.2 |

Fuente: Ragin (2008) con adaptación propia.

2. En dicha tabla, se muestra una matriz del conjunto de *datos difusos* para el caso hipotético de cómo un grupo de empresas innovadoras ubicadas dentro de un clúster de innovación tecnológico de manufactura, presenta su toma de decisiones dentro de un clúster de innovación, siendo de principal interés, el grado de pertenencia en el conjunto de empresas innovadoras *calidad de votación por liderazgo tecnológico,* **W** y se asumen estables en un periodo de tiempo largo (**20** años). Este análisis se centra en las condiciones ligadas más que nada a las votaciones de las empresas innovadoras clúster con bajo nivel de calidad de votación. Las *condiciones causales* usadas en el ejemplo, son:
   - **a.** Grado de pertenencia en el conjunto de empresa innovadoras clúster con sólidas *fortalezas en resultados de sus consejos de administración* **U.**
   - **b.** Grado de pertenencia en el conjunto empresas innovadoras clúster con empleo de un *alto grado tecnológico de manufactura,* **M.**
   - **c.** Grado de pertenencia en el conjunto de *empresas innovadoras clúster por su nivel de in luencia,* **A.**
   - **d.** Grado de pertenencia en el conjunto de *empresas innovadoras clúster con diferencia en sus ingresos por innovación,* **I.**
3. En el conjunto de empresa innovadoras clúster con sólidas *fortalezas en resultados de sus consejos de administración* **U** y con empleo de un *alto grado tecnológico de manufactura* **M**, tienden a tener una fuerte *calidad de votación por liderazgo tecnológico* **W**, mientras las que tienen *nivel de in luencia* **A** y *diferencia de ingresos por innovación* **I**, tienden a socavarla.
4. Los puntajes difusos de pertenencia, reflejan caracterizaciones generales de las empresas, basadas en las observaciones y la experiencia del investigador en su categorización.
5. Es importante señalar, que el enfoque esbozado en este capítulo hace énfasis en el uso de la *tabla de verdad* del conjunto de *datos difusos*, como dispositivo analítico clave.
6. Una ventaja adicional de este enfoque, es que es más transparente y, por lo tanto, el investigador tiene un control más directo sobre el análisis.

7. Este tipo de control es central para la práctica de la investigación orientada a casos. El enlace, entre conjunto de *datos difusos* con las *tablas de verdad*, tiene tres pilares principales. Ver **Tabla 4.14.**

**Tabla 4.14. Pilares principales que relacionan el conjunto de datos difusos (*fuzzy-sets*) con las tablas de verdad**

| Pilar | Descripción |
|---|---|
| **1** | La *correspondencia directa* que existe entre las filas de una *tabla de verdad* y las *esquinas del espacio vectorial* definidas por el conjunto de *datos difusos* (*fuzzy-sets*) de las *condiciones causales* **(Ragin, 2000).** |
| **2** | La *evaluación de la distribución de casos* a través de las diferentes combinaciones lógicas de las *condiciones causales* (es decir, la distribución de casos a través de sectores del espacio vectorial, definidos por las *condiciones causales* del conjunto de *datos difusos* (*fuzzy-sets*). |
| **3** | La *evaluación de la consistencia de la evidencia.* para cada combinación causal con el argumento de que es un subconjunto del resultado. Como se explicó en el capítulo 1, la relación de subconjunto es importante, porque señala una conexión explícita entre una combinación de *condiciones causales* y un resultado. |

Fuente: Ragin (2008) con adaptación propia.

8. Una vez que estos tres pilares estén alineados, es posible construir una *tabla de verdad* que resuma los resultados de análisis del conjunto de *datos difusos* y posteriormente, analizar esta *tabla de verdad.* En efecto, el análisis de la *tabla de verdad* sintetiza los resultados de múltiples análisis de conjuntos de *datos difusos*.

## Correspondencia entre las esquinas del espacio vectorial y filas de la tabla de verdad

Un espacio vectorial multidimensional, construído a partir de un conjunto de *datos difusos* de *condiciones causales*, tienen $2^k$ esquinas, así como una *tabla de verdad* nítida con $2^k$ filas (donde $k$ es el número de *condiciones causales*). Existe una correspondencia directa entre las combinaciones

causales, filas de *tablas de verdad* y las *esquinas de espacio vectorial* **(Ragin, 2000)**. Por ejemplo, una *tabla de verdad* simple, con dos conjuntos de *condiciones causales* nítidas:

1. Tiene cuatro filas: **(00), (01), (10) y (11**); un espacio vectorial formado a partir de dos *condiciones causales* de conjunto difuso tiene cuatro esquinas: **(0,0), (0,1), (1,0) y (1,1).**
2. En el análisis del conjunto de datos nítidos, los casos se clasifican en la *tabla de verdad*, dentro de sus filas, según sus combinaciones específicas de puntación de presencia / ausencia, en las *condiciones causales*. Por lo tanto, cada caso se asigna a una fila única, y cada fila abarca un subconjunto único de los casos incluídos en el estudio.
3. Con conjuntos de *datos difusos*,sin embargo, los casos pueden variar en grados de pertenencia, en las diferentes *esquinas del espacio vectorial* y por tanto, variando los grados de pertenencia en las filas de la *tabla de verdad* correspondiente.
4. El grado de pertenencia de un caso, en una esquina del espacio vectorial del conjunto de *datos difusos*, está determinado por sus puntajes de pertenencia.
5. Por ejemplo, un caso con una pertenencia de **0.7** en el conjunto **B** y una pertenencia de **0.9** en el conjunto **C,** tiene una pertenencia de **0.7** en **(1,1)** de la esquina **(B * C)** del espacio vectorial formado por estos dos conjuntos difusos (el mínimo de **0.7** y **0.9** es **0.7**).
6. Este mismo caso, tiene una pertenencia de **0.1** en **(0,0)** de esquina **(~ B * ~ C)** del espacio vectorial. La pertenencia de este caso en (**~ B * ~ C**)**= min (~ B , ~ C)= min (1 - 0.7, 1 - 0.9)= min (0.2, 0.1)= 0.1,** donde **~** señala la negación del conjunto de *datos difusos*, que corresponde a la ausencia usando conjuntos de relaciones nítidas y el símbolo **( * )** que denota un conjunto de intersecciones (lógica **AND**).
7. A pesar de esta característica de los conjuntos de *datos difusos*, las *tablas de verdad* todavía se pueden usar, para ayudar a su análisis.
8. Específicamente, el investigador puede usar las filas de las *tablas de verdad*, como especificaciones de las *esquinas del espacio vectorial* y

usar la *tabla de verdad* para resumir, las declaraciones sobre las características de la combinación de causalidad representada por cada esquina.

Por ejemplo, el investigador podría calcular el grado de pertenencia en la esquina **(~ A * ~ I * M * U)** del espacio vectorial (la combinación de una baja pertenencia del: *nivel in luencia* **A,** con *baja diferencia de ingresos* **I**, y alta pertenencia del: grado *tecnológico de manufacturera* **M** y la *fortaleza de resultados del consejo de administración,* **U**) y evaluar, si el grado de pertenencia en esta esquina del espacio vectorial es un subcon-junto consistente de grado de pertenencia en el resultado (*calidad de votación*, **W**). Así:

1. El investigador usaría las técnicas esbozadas en este capítulo, para evaluar el subconjunto de *datos difusos* y usar los puntajes de pertenencia para todos los casos en esta evaluación, no solo aquellos con puntuación sólida en su pertenencia.
2. Si el grado de pertenencia en esta esquina del espacio vectorial (que corresponde a una de las dieciséis combinaciones de *condiciones causales*) es consistentemente menor o igual al grado de pertenencia en el resultado en todos los casos, entonces el investigador puede concluir, que la pertenencia a la combinación causal es un subconjunto de la pertenencia a el resultado.
3. El investigador posteriormente agregaría información sobre los resultados de esta evaluación a la fila de la *tabla de verdad* correspondiente, a la esquina relevante del espacio vectorial. De esta manera, toda la *tabla de verdad*, se puede utilizar para resumir los resultados del $\mathbf{2}^{k}$ análisis de conjuntos de *datos difusos*.
4. Así, en la traducción de conjuntos difusos a *tablas de verdad*, esta representa las declaraciones sobre las *esquinas del espacio vectorial* formadas por las *condiciones causales* del conjunto difuso. De esta forma, se destacan dos aspectos sobre las *esquinas del espacio vectorial*, que son especialmente importantes:

**a.** El número de casos con una fuerte pertenencia en cada esquina (es decir, en cada combinación de *condiciones causales*) y

**b.** La *consistencia* de la evidencia empírica para cada esquina, con el argumento que el grado de pertenencia en cada esquina del espacio vectorial, es un subconjunto de grado de pertenencia en el resultado.

## Distribución de casos a través de las combinaciones causales

La distribución de casos, a través de las combinaciones causales es fácil de evaluar cuando las *condiciones causales* se representan:

1. Con conjuntos de relaciones nítidas, porque es un forma sencilla de construir una *tabla de verdad* a partir de dichos datos y de examinar el número de casos nítidamente ordenados, dentro de cada fila.
2. Cuando las *condiciones causales* son conjuntos de *datos difusos,* sin embargo, este análisis es menos directo porque cada caso puede tener una pertenencia parcial en cada fila de la *tabla de verdad* (es decir, en cada esquina del espacio vectorial).
3. Aún así, es importante evaluar la distribución de las puntuaciones de pertenencia de los casos entre combinaciones causales en análisis de conjuntos de *datos difusos* porque, algunas combinaciones causales pueden ser empíricamente triviales.
4. En otras palabras, si la mayoría de los casos tienen muy baja o nula pertenencia en una combinación causal, entonces no tiene sentido evaluar qué combinación produce el resultado. La base empírica para tal evaluación, sería demasiado débil para justificarlo. Algunas *esquinas del espacio vectorial* pueden tener muchos casos con pertenencia fuerte; otras esquinas pueden tener solo unos pocos casos, con algún grado de pertenencia.
5. Al construir una *tabla de verdad* a partir de conjuntos de relaciones difusas, es importante tomar en cuenta estas diferencias. La **Tabla 4.15,** muestra la distribución de los puntajes de pertenencia de doce empresas

innovadoras clúster a través de las dieciséis combinaciones lógicamente posibles de las cuatro *condiciones causales* **(A, M, I, U)** que se exponen en la **Tabla 4.15.**

6. Las letras minúsculas en esta tabla, indican la negación de una condición **(a = ~ A).** En esencia, la tabla lista las dieciséis *esquinas del espacio vectorial* de cuatro dimensiones, que está formada por los cuatro conjuntos de *datos difusos* y muestra también el grado de pertenencia de cada caso, en cada esquina.
7. Esta tabla, también muestra una importante propiedad de la intersección de los conjuntos de *datos difusos*, a saber, que cada caso puede tener (como máximo) solo un puntaje de pertenencia mayor que **0.5** en las posibles combinaciones formadas, a partir de un conjunto dado de *condiciones causales*. Tenga en cuenta, que si un caso tiene **0.5** de pertenencia en una *condición causal*, entonces su máximo puntaje de pertenencia, en cualquier combinación causal que incluya esa condición, es solo **0.5**. Así, cualquier caso con un código de **0.5** en una *condición causal*, no será más cercano a ninguna esquina del espacio vectorial, definido por las *condiciones causales*. Por lo tanto, es una buena práctica evitar, si es posible, usar el puntaje de pertenencia de **0.5**, que indica la máxima ambigüedad, al evaluar el grado de pertenencia en una *condición causal*.
8. Un puntaje de pertenencia mayor a **0.5**, en una combinación causal, señala que un caso está más dentro que fuera, de la combinación causal en cuestión. Un puntaje mayor a **0.5**, también indica a qué esquina del espacio vectorial multidimensional, un caso dado se encuentra más cercano. Esta propiedad de los conjuntos difusos, permiten a los investigadores determinar cuántos casos están cerca a cada esquina del espacio vectorial, según su grado de pertenencia.

**Tabla 4.15. Evaluando la distribución de casos entre las combinaciones de *condiciones causales***

| EMPRESA | Columnas | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| | a*i*m*u | a*i*m*U | a*i*M*u | a*i*M*U | a*I*m*u | a*I*m*U | a*I*M*u | a*I*M*U | A*i*m*u | A*i*m*U | A*i*M*u | A*i*M*U | A*I*m*u | A*I*m*U | A*I*M*u | A*I*M*U |
| A | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.6 | 0.4 | 0.4 |
| B | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| C | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 |
| D | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 |
| E | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.2 | 0.2 | 0.2 | 0.6 | 0.4 | 0.4 | 0.4 |
| F | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| G | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 |
| H | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.6 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 |
| I | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.6 | 0.2 | 0.4 | 0.2 | 0.4 |
| J | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.6 | 0.0 | 0.2 | 0.0 | 0.4 |
| K | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.6 |
| L | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.4 | 0.2 |
| S | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 2 | 0 | 2 | 3 | 1 | 0 | 1 |

Notas: S. Score > 0.5

Fuente: Elaboración propia.

9. La última fila de la **Tabla 4.15** muestra el número de casos mayores a la puntuación de **0.5** de pertenencia en cada esquina. La tarea clave en esta fase del análisis, es establecer un número de umbral de casos, es decir, desarrollar una regla para clasificar algunas combinaciones de condiciones como relevantes y otras como restantes basados en el número de casos con más de **0.5** de pertenencia en cada combinación. Una fila restante o *residual* en el análisis cualitativo comparativo por conjuntos difuso **fsQCA**, es una combinación lógicamente posible de condiciones, que carecen de instancias de valor empírico, ya sea porque el investigador tiene información inadecuada sobre tales casos o porque los casos simplemente no existen.
10. La regla establecida por el investigador, debe reflejar la naturaleza de la evidencia y el carácter del estudio. Las consideraciones importantes, incluyen: el número total de casos, el número de *condiciones causales*, el grado de familiaridad del investigador con cada caso, el grado de precisión que es posible hacer en la calibración de conjuntos de *datos difusos*, el alcance de la medición y asignación de error, si el investigador está interesado en patrones gruesos versus graduados a detalle en los resultados, y así sucesivamente.
11. El conjunto de datos utilizado en esta demostración, se compone de doce casos y dieciséis combinaciones lógicamente posibles de condiciones. En esta situación, un umbral de frecuencia razonable es un caso. Así, los ocho combinaciones de condiciones que carecen de un solo caso con más de **0.5** de pertenencia, se tratan como residuos o restantes en el análisis que sigue, debido a que no hay instancias empíricas fuertes de ninguna de estas combinaciones.
12. Las combinaciones causales, con al menos un caso mayor que **0.5** de pertenencia, se retienen para un examen más detallado. El tratamiento de residuos o restantes es muy importante en el tratamiento de **fsQCA**.
13. Cuando el número total de casos, incluidos en un análisis es grande (por ejemplo, cientos), es importante establecer un umbral de frecuencia *(frequency threshold)* para la relevancia o viabilidad de las combinaciones causales.

**14.** En tales análisis, algunas *esquinas de espacio vectorial*, pueden tener varios casos con más de **0.5** de pertenencia, debido a errores de medida o codificación. Sería prudente, por lo tanto, tratar las combinaciones causales de baja frecuencia iguales a las que carecen de fuerza como instancias empíricas por completo (es decir, las mismas pero con frecuencia = **0**).

**15.** Cuando el número total de casos en un estudio es grande, el problema no es qué combinaciones causales tienen instancias (es decir, al menos un caso mayor de **0.5** de pertenencia), sino qué combinaciones causales tienen suficientes instancias para garantizar la realización de una evaluación de la relación de subconjunto con el resultado. Por ejemplo, un investigador podría usar una frecuencia de umbral de al menos diez casos, con más de **0.5** de pertenencia.

## Evaluando la consistencia de los subconjuntos de datos difusos (*fuzzy-sets*)

Una vez que se han identificado las combinaciones causales empíricamente relevantes basados en la distribución de casos, el siguiente paso, es evaluar la *consistencia* de cada combinación con el marco teórico de la investigación en cuestión. Se plantean preguntas como ¿qué combinaciones causales, son subconjuntos del resultado?, ¿qué diferentes combinaciones de condiciones causales están vinculadas?, ¿existe una fuerte puntuación de pertenencia en el resultado? , por lo que:

1. La medición de la *consistencia* del marco teórico, utilizada para evaluar el vínculo entre las combinaciones causales y sus resultados utiliza la fórmula:

$$\textbf{Consistencia} \; (X_i <= Y_i) = \Sigma \, [\min (X_i , Y_i)] \, / \, \Sigma \, (X_i)$$

2. Donde **min** indica la selección más baja de dos valores; $X_i$ es el grado de pertenencia en una combinación causal; $Y_i$ es el grado de pertenencia en el resultado **(Kosko, 1993; Smithson y Verkuilen 2006)**.

3. Cuando los valores $X_i$ son todos menores o iguales a sus valores de $Y_i$ correspondientes, el puntaje de *consistencia* es **1.00**; cuando solo unos pocos no cumplen la condición, el puntaje es ligeramente menor a **1.00**; cuando son muchos los que no la cumplen, hay puntuaciones inconsistentes con algunos valores de $X_i$ muy superiores a sus valores de correspondientes, por lo que la *consistencia* puede caer por debajo de **0.5**.
4. La **Tabla 4.16,** muestra el grado en que las ocho combinaciones causales relevantes (aquellos, con al menos un caso con más de **0.5** de pertenencia) son subconjuntos consistentes del resultado (*calidad de votación*, **W**) usando datos sobre las doce empresas innovadoras clúster, para cada evaluación.

**Tabla 4.16. Evaluando la consistencia de combinaciones cusales con la relación de subconjuntos de datos difusos (*fuzzy-subsets*)**

| Empresas innovadoras Clúster | Columnas | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | Resultado calidad de votación por liderazgo tecnológico | Consistencia | Nivel de influencia | Diferencia de ingresos por innovación | Grado tecnológico de manufactura | Fortaleza resultados consejo de administración |
| | W | C | A | I | M | U |
| **1** | 1 | 1.00 | 1 | 0 | 0 | 0 |
| **2** | 1 | 1.00 | 1 | 1 | 0 | 0 |
| **3** | 0 | 0.87 | 0 | 1 | 1 | 1 |
| **4** | 0 | 0.84 | 1 | 1 | 0 | 1 |
| **5** | 0 | 0.82 | 0 | 1 | 0 | 1 |
| **6** | 0 | 0.79 | 1 | 0 | 0 | 1 |
| **7** | 0 | 0.78 | 1 | 1 | 1 | 1 |
| **8** | 0 | 0.72 | 1 | 0 | 1 | 1 |

Fuente: Ragin (2008) con adaptación propia.

5. Las puntuaciones de *consistencia* reportados se basan en la fórmula presentada anteriormente.
6. Para facilitar la interpretación, las combinaciones causales han sido ordenados en orden descendente según su *consistencia* puntuaciones. Es importante recordar que la **Tabla 4.16**, en efecto, presenta declaraciones resumidas sobre las *esquinas del espacio vectorial*, definido por las cuatro *condiciones causales* del conjunto de *datos difusos*, utilizadas en este análisis.
7. Cada fila, esencialmente responde a la pregunta: ¿es el grado de pertenencia en cierta esquina del espacio vectorial, un subconjunto consistente de grado de pertenencia en el resultado? El análisis de la evidencia en la **Tabla 4.16** es, por lo tanto, una síntesis lógica de las declaraciones sobre las *esquinas del espacio vectorial*.

## Analizando la tabla de verdad

Las determinaciones clave, que deben hacerse, son:

1. Es la puntuación de *consistencia* que se utilizará como valor de corte para determinar qué combinaciones causales pasan la *consistencia* difusa del marco teórico y cuáles otras no.
2. Las *combinaciones causales*, con puntajes de *consistencia* por encima del valor de corte, se designaron como subconjuntos de *datos difusos*, del resultado y se codifican como **1**.
3. Los que están por debajo del valor de corte no son considerados como subconjuntos de *datos difusos*, o borrosos y se codifican como **0**. Filas que no alcanzan el umbral de frecuencia seleccionado por el investigador (según el número de casos con más de **0.5** de pertenencia) se tratan como filas *residuales*. La designación de tales filas, como *residuales*, se justifica porque la evidencia empírica relevante para estas combinaciones, no es lo suficientemente sustancial como para permitir una evaluación de la consistencia del marco teórico.

4. En efecto, las combinaciones causales, que son subconjuntos de *datos difusos* del resultado, delinean los tipos de casos en los que el resultado es encontrado (por ejemplo, los tipos empresas innovadoras clúster que tienen un cierto tipo de *calidad de votación por liderazgo tecnológico*, **W**).
5. Una simple inspección de los valores de *consistencia* en la **Tabla 4.16,** revela una considerable brecha en los puntajes de *consistencia* entre la *segunda y la tercera combinación causal*; el grado de *consistencia* con la relación del subconjunto cae de **1.00** (*consistencia* perfecta) a **0.87**.
6. Esta brecha, proporciona una base fácil para diferenciar combinaciones causales consistentes de combinaciones inconsistentes, como se muestra en la *columna 1* de la **Tabla 4.16**, resultado de *calidad de votación por liderazgo tecnológico*,**W.**
7. En la mayoría de los análisis de este tipo, sin embargo, el valor de corte de la *consistencia*, será menor que **1.0** ; por cierto, una *consistencia* de marco teórico perfecto, no es común con datos difusos.

**Ragin (2000)**, demuestra cómo incorporar criterios probabilísticos dentro de la evaluación de la *consistencia* de subconjuntos de relaciones, y estos mismos criterios, puedan modificarse para su uso aquí, de forma que:

1. La prueba probabilística, requiere un valor de referencia (por ejemplo, *consistencia* **0.75**) y un alfa (por ejemplo, significancia de **0.05**). En el interés de permanecer cerca de la evidencia, es a menudo útil, simplemente ordenar las puntuaciones de *consistencia* en orden descendente, como en la **Tabla 4.16,** y observe si ocurre una brecha sustancial, en la parte superior de las puntuaciones de los rangos de *consistencia*.
2. Con la diferencia entre las filas dos y tres de la **Tabla 4.16,** está claro que usar criterios probabilísticos para ayudar a la selección de un valor de corte, simplemente ofuscaría lo obvio.
3. En general, el valor de corte no debe ser inferior a **0.75**; un valor de corte ≥ **0.85**, es lo recomendado especialmente para datos de macro nivel. Mientras la medición de *consistencia* puede variar de **0.0** a **1.0**, las

puntuaciones entre **0.0** y **0.75** indican la existencia de una in*consistencia* sustancial.

4. La *tabla de verdad* resultante de estos procedimientos está contenida en la misma **Tabla 4.16.** Las *columnas del 2 al 6* muestran las codificaciones de las *condiciones causales*; la *columna 1* muestra el resultado de datos nítidos (*consistente* vs. *no consistente*) adjunto a cada fila de la *tabla de verdad* (*esquina del espacio vectorial*).
5. Las ocho combinaciones de *condiciones causales* que no figuran en la **Tabla 4.16** son *residuales*. Los resultados del análisis de esta *tabla de verdad*, con las combinaciones residuales, definidas como falsas (es decir, sin supuestas simplificaciones permitidas) se muestra:

**(A * ~ M * ~ U) < = W**

6. Donde los grados de pertenencia de las empresas innovadoras clúster, son: **W.** *Calidad de votación por liderazgo tecnológico;* **A.** Nivel de influencia; **M.** *Alto grado tecnológico de manufactura;* **U.** Fortalezas en resultados de sus consejos de administración; **(~)** indica conjuntos difusos negados y ≤ señala al subconjunto relación.
7. La *consistencia* marco teórico de la investigación de esta solución, es **1.0;** la *cobertura* del resultado (*votación de clase*, **W** débil) es **0.636** (ver cálculo de cobertura más adelante)**.** Los resultados indican que la calidad *de votación por liderazgo tecnológico* **W** débil, ocurre en empresas innovadoras del clúster, donde *el nivel de influencia* **A,** se combina con una pertenencia débil en el conjunto de empresas innovadoras clúster con *alto grado tecnológico de manufactura* **M** y pertenencia débil, en el conjunto de empresas innovadoras clúster, con fortalezas en resultados de sus consejos de administración, **U.**
8. La inspección de la **Tabla 4.13**, revela que las mejores instancias de esta combinación de condiciones son las empresas innovadoras clúster: **(L, D, E y H ).**
9. Cuando los *supuestos simplificadores* se extraen del grupo los de ocho *residuales* (por **fsQCA**), resulta una solución más *parsimoniosa*:

**(~ U) < = (W)**

**10.** Considerando nuevamente la **Tabla 4.13.**

**Tabla 4.13. Conjunto de datos difusos (*fuzzy-sets*) para el caso de decisión de clúster de innovación**

| Empresas innovadoras Clúster | Columnas | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | Calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado tecnológico de manufactura | Fortaleza resultados consejo de administración |
| | W | A | I | M | U |
| A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 |
| C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 |
| D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 |
| E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 |
| F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 |
| G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 |
| H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 |
| I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 |
| J | 0.0 | 0.8 | 0.4 | 0.8 | 1.0 |
| K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 |
| L | 1.0 | 1.0 | 0.8 | 0.4 | 0.2 |

Fuente: Elaboración propia.

Según estos resultados, la *calidad de votación por liderazgo tecnológico* **W** débil, tiene una fuente única motivada por las *fortalezas en resultados de sus consejos de administración* **(~ U)** nulas**.** La *consistencia* del marco teórico de la investigación de esta solución, es **1.0**; la *cobertura* del resultado es **0.727** (ver cálculo de cobertura más adelante).

**11.** La **Tabla 4.13** las empresas innovadoras del clúster, con los puntajes de pertenencia más bajos en *fortalezas en resultados de sus consejos de administración* **U** son **L** y **D.** Esta solución de la *tabla de verdad,*

depende en seis supuestos simplificadores **(Ragin 1987, 2000)** que describe combinaciones de condiciones de causalidad no observadas.

**12.** Los **6** supuestos se extraen de las **8** combinaciones. de *condiciones causales* que carecen de instancias empíricas fuertes (es decir, aquellos que carecen de casos con puntajes de pertenencia mayores a **0.5**, en la última fila de la **Tabla 4.15**).

**Ragin y Sonnett (2004)**, demuestran cómo usar estas dos soluciones para realizar análisis *contrafactuales*. La primera solución maximiza la *complejidad*; el segundo maximiza la *parsimonia*.

**13.** En esta coyuntura, es importante señalar una propiedad de los conjuntos de *datos difusos*, que los distingue claramente de los conjuntos de relaciones nítidos.

**14.** En pocas palabras, con conjuntos de *datos difusos* es matemáticamente posible, para una *condición causal* o combinación causal, ser un subconjunto de ambos resultados (por ejemplo, *votación de clase,* **W** débil) y la negación del resultado (por ejemplo, *calidad de votación por liderazgo tecnológico* **W**, no débil). El resultado es matemáticamente posible, porque el grado de pertenencia en un *condición causal* o combinación causal (por ejemplo, una puntuación de **0.3**), puede ser menor que el resultado (por ejemplo, **0.6**) y menos que la negación del resultado **(1.0 - 0.6 = 0.4).**

**15.** También es posible que una condición o combinación causal sea inconsistente con el resultado y su negación, al exceder ambos (por ejemplo, el puntaje de combinación causal = **0.8**; la puntuación de pertenencia del resultado= **0.7**; la negación de la puntuación de pertenencia de resultado = **0.3**).

**16.** El punto importante, es que con conjuntos de *datos difusos*, no existe una razón matemática para esperar puntajes de *consistencia* calculados para la negación de un resultado que esté perfectamente correlacionado negativamente, con puntajes de *consistencia* para el resultado original, como los que se encuentran en los análisis de conjuntos de datos nítidos.

## Tabla 4.15. Evaluando la distribución de casos entre las combinaciones de *condiciones causales*

| EMPRESA | Columnas | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| | a*i*m*u | a*i*m*U | a*i*M*u | a*i*M*U | a*I*m*u | a*I*m*U | a*I*M*u | a*I*M*U | A*i*m*u | A*i*m*U | A*i*M*u | A*i*M*U | A*I*m*u | A*I*m*U | A*I*M*u | A*I*M*U |
| A | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.6 | 0.4 | 0.4 |
| B | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| C | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 |
| D | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.2 | 0.2 |
| E | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.2 | 0.2 | 0.2 | 0.6 | 0.4 | 0.4 | 0.4 |
| F | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| G | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.6 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 |
| H | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.6 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 |
| I | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.4 | 0.2 | 0.6 | 0.2 | 0.4 | 0.2 | 0.4 |
| J | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.2 | 0.0 | 0.6 | 0.0 | 0.2 | 0.0 | 0.4 |
| K | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.2 | 0.2 | 0.4 | 0.6 |
| L | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.2 | 0.2 | 0.2 | 0.2 | 0.6 | 0.2 | 0.4 | 0.2 |
| **Score>0.5** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 2 | 0 | 2 | 3 | 1 | 0 | 1 |

Fuente: Elaboración propia.

**17.** Esta propiedad, de los conjuntos de *datos difusos*, proporciona una razón adicional para realizar un análisis de la negación del resultado (por ejemplo, *calidad de votación por liderazgo tecnológico*, **W,** no débil) por separado del análisis del resultado (por ejemplo, *calidad de votación por liderazgo tecnológico* **W**, débil).

**18.** El llevar estos dos análisis por separado, permite la asimetría, entre los resultados del análisis de las causas de un resultado y el análisis de la causas de su negación.

**19.** El análisis separado de una salida y su negación, es también una práctica utilizada en el análisis cualitativo comparativo nítido **csQSA**.

**20.** Desde el punto de vista de los métodos correlacionales, esta propiedad del marco teórico de la investigación es desconcertante. Desde el punto de vista de la teoría social, sin embargo, no lo es.

**21.** Considere la solución utilizando supuestos simplificadores presentados arriba. Este análisis muestra, que la *calidad de votación por liderazgo tecnológico* **W**, persistentemente en niveles bajos, se encuentran en empresas innovadoras del clúster con *fortaleza resultados consejo de administración* **U**, débiles. En efecto, este análisis revela un impedimento principal para la *calidad de votación por liderazgo tecnológico,* **W.** La pregunta de ¿cuáles *condiciones causales* son *impedimentos* para la calidad de *votación por liderazgo tecnológico,* **W**?, no es la misma a la pregunta de ¿cuáles *condiciones causales* son productivas para la calidad de votación?

**22.** Las respuestas a estas dos preguntas podrían diferir fácilmente. Así, la asimetría del análisis marco teórico de la investigación, se alinea con las expectativas teóricas de la causalidad asimétrica.

## Conveniencia de los conjuntos de datos difusos (*fuzzy-sets*) de no hacerlos dicotómicos

La construcción de una *tabla de verdad* a partir de *datos difusos*, como se acaba de bosquejar, es cognitivamente exigente. Implica dos análisis:

**1.** La evaluación de la distribución de casos entre combinaciones causales y
**2.** La evaluación del grado de *consistencia* de cada combinación causal, con la relación de subconjunto frente al resultado.

Además, ambas evaluaciones implican la selección de valores de corte, que pueden parecer arbitrarios. En realidad, el rango de valores de corte plausibles es relativamente estrecho, y el rango puede reducirse aún más, cuando los investigadores están familiarizados con sus casos y con literatura de marco teórico sustantivo.

Debido a lo anterior, cabe preguntarse ¿por qué no simplemente recodificar el conjunto de *datos difusos*, en conjuntos de datos nítidos y realizar un análisis convencional en este modo, utilizando los puntajes de pertenencia dicotomizados? Después de todo, un puntaje difuso de **0.5**, diferencia los casos que son más dentro vs. más fuera, de un conjunto de *datos difusos* dado. El uso del valor de cruce o traslape, para crear un conjunto de datos nítidos, parece ser una extensión sencilla del enfoque.

La mejor manera de evaluar la viabilidad de esta opción, es simplemente volver a analizar los datos del conjunto difuso presentados en la **Tabla 4.13**, convirtiendo primero el conjuntos difusos a conjuntos nítidos. La **Tabla 4.17** muestra los datos nítidos que resultan a partir de la aplicación de la regla de punto de cruce (que dicotomiza los *datos difusos* en el puntaje de pertenencia de **0.5**) a los datos de conjunto difuso presentados en **Tabla 4.13**.

Tabla 4.17. Conjunto de datos nítidos (*crisp-sets*)

| Empresas innovadoras Clúster | Columnas | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | Resultado calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado tecnológico de manufactura | Fortaleza resultados consejo de administración |
| | W | A | I | M | U |
| **A** | 1 | 1 | 1 | 0 | 1 |
| **B** | 1 | 1 | 0 | 0 | 1 |
| **C** | 0 | 1 | 0 | 0 | 1 |
| **D** | 1 | 1 | 1 | 0 | 0 |
| **E** | 1 | 1 | 1 | 0 | 0 |
| **F** | 1 | 0 | 1 | 1 | 1 |
| **G** | 1 | 0 | 1 | 0 | 1 |
| **H** | 1 | 1 | 0 | 0 | 0 |
| **I** | 0 | 1 | 0 | 1 | 1 |
| **J** | 0 | 1 | 0 | 1 | 1 |
| **K** | 0 | 1 | 1 | 1 | 1 |
| **L** | 1 | 1 | 1 | 0 | 0 |

Fuente: Elaboración propia.

Claramente, se obtiene una ganancia en simplicidad al comparar **Tablas 4.13 y 4.17**. Los casos en la **Tabla 4.17** están o completamente dentro **(1)** o completamente fuera **(0)** de los conjuntos relevantes; en la **Tabla 4.13**, sus pertenencias se califican. Sin embargo, algunos costos están involucrados en la obtención de esa simplicidad. Observe, por ejemplo, que el conjunto de datos, ahora incluye una contradicción (casos idénticos con resultados contrarios). Las empresas innovadoras clúster **(B y C)** tienen los mismos puntajes en las cuatro *condiciones causales* nítidas (**A, I, M y U**), pero tener diferentes puntuaciones en el resultado **W**. En un análisis del conjunto de datos nítidos convencional, es necesario abordar esta contradicción de alguna manera antes que el análisis pueda continuar.

Con una combinación causal contradictoria y ocho *residuales*, hay varias formas de analizar la evidencia en la **Tabla 4.17**. En un esfuerzo para que coincida con la segunda solución de conjunto *datos difusos* lo más cerca posible, la contradicción se establece en falso y los ocho restantes se usan como combinaciones neutrales, lo que los hace disponibles para su uso como supuestos de simplificación. Se puede asignar una fila de neutral **1 o 0** en el resultado por **fsQCA**, dependiendo de qué tarea tenga más solución *parsimoniosa*. Los resultados del análisis de conjunto nítido muestran:

**(~ U ) + (~ A ) + (I * ~ M) < = W**

1. Donde:
   **W.** Calidad de votación por liderazgo tecnológico;
   **A.** Nivel de influencia;
   **I.** Diferencia de ingresos por innovación;
   **M.** Grado tecnológico de manufactura;
   **U.** Fortaleza resultados consejo de administración;
   **(~).** señales de negación; **(*)** Indica **AND** lógico (o combinaciones de condiciones), **(+)** indica **OR** lógico (o *condiciones causales* alternativas o combinaciones causales alternativas), y **(< =)** indica la relación del subconjunto. Los resultados indican, que hay tres bases alternativas para la calidad de votación por liderazgo tecnológico **W,** débil:

   **a.** Baja fortaleza resultados consejo de administración **(~U)**
   **b.** Bajo nivel de influencia **(~A)**
   **c.** Una combinación de diferencia de ingresos y menor grado tecnológico de manufactura: **(I * ~ M).**

2. Esta solución depende de una serie de supuestos simplificadores (no examinados por el momento aquí) porque ocho de las dieciséis filas de la *tabla de verdad,* carecen de casos.
3. La diferencia clave entre la solución de conjuntos nítidos y la segunda solución de conjuntos difusos, es decir, la solución más *parsimoniosa,* es que la solución de conjuntos nítidos, agrega dos términos nuevos:

**(~ A)** (*bajo nivel de influencia*) y **(I * ~ M)** (sustancial desigualdad de ingresos combinada con menor grado tecnológico de manufactura).

4. Por lo tanto, la solución de conjuntos nítidos es más compleja y más inclusivo que la segunda solución del conjunto de datos difuso.
5. Los dos nuevos términos, **(~ A)** y **(I * ~ M),** aparecen en la solución de conjuntos de datos nítidos debido a su estándar más bajo de *consistencia* del marco teórico de la investigación.
6. La *consistencia* perfecta del marco teórico de la investigación, se logra con conjuntos de *datos difusos* cuando todos los casos se encuentran por encima de la diagonal principal de la gráfica de dispersión difusa.
7. Con conjuntos de datos nítidos, sin embargo, la *consistencia* perfecta del marco teórico, es mucho más fácil de lograr. Mientras no aparezcan casos en el cuarto cuadrante de la gráfica de dispersión difusa (el cuadrante inferior derecho), posteriormente el conjunto trazado en el eje horizontal, puede describirse como un subconjunto, del conjunto trazado en el eje vertical. El cuadrante inferior derecho, de un diagrama de dispersión de dos conjuntos difusos, es un subconjunto de y solo la mitad del tamaño del triángulo debajo de la diagonal principal. Este estándar de menor *consistencia* define más casos y por lo tanto combinaciones más causales como consistentes.

Por ejemplo, tanto las empresas innovadoras **C** e **I** que obtuvieron un puntaje de **0.4** en **(~ A)** (*bajo nivel de in luencia*, en términos de solución por conjuntos nítidos) y **0.2** en **W** (*calidad de votación por liderazgo tecnológico débil*), por lo que:

1. En el análisis de conjuntos de *datos difusos*, estos casos se consideran como inconsistentes porque sus puntajes en la *condición causal*, excede sus puntajes en el resultado.
2. Las *inconsistencias* socavan directamente el argumento de que **(~ A)** es un subconjunto de **W.**
3. Sin embargo, desde una perspectiva del conjunto de datos nítidos, estos casos son enteramente consistentes porque muestran una ausencia de

influencia **A** (su puntaje de **0.4**, se recodifican a **0**) y la ausencia de una calidad de votación por liderazgo tecnológico **W** o débil (con puntaje de **0.2,** también se recodifica a **0**).

4. Un patrón similar se tiene con **(I * ~ M):** algunos de los análisis de los casos definidos como consistentes, en los conjuntos de datos nítidos, son inconsistentes en el análisis de conjuntos de *datos difusos*. Por lo tanto, los términos causales adicionales que aparecen en la solución de conjunto nítido se deben a su estándar de menor *consistencia*.
5. Dados estos resultados, parece que la práctica de dicotomizar conjuntos de *datos difusos* (y, por implicación, escalas de intervalo y razón) para crear conjuntos de datos nítidos para análisis convencionales de conjuntos de datos nítidos, no son una opción atractiva.
6. Los investigadores deben usar los procedimientos descritos para calibrar los conjuntos de *datos difusos* de escalas de intervalo y razón y mejor realizar análisis de conjuntos de *datos difusos*.
7. Usar conjuntos de *datos difusos* también, es preferible a usar conjuntos multivalor (*multivalue-sets*, **Cronqvist, 2004**) porque estos tienden a exacerbar el problema de diversidad limitada (**Rihoux y Ragin, 2009**).
8. De manera más general, estos resultados indican que si los investigadores representan sus *condiciones causales* y resultados como conjuntos de *datos difusos*, ellos no deberán, por lo tanto, usar conjuntos de datos nítidos, los cuales se sugiere, deben reservarse para fenómenos de naturaleza categórica.

El objetivo de **fsQCA** es ayudar a la interpretación causal, a la par con el conocimiento de casos. El objetivo práctico de las técnicas presentadas. en este capítulo, y de **fsQCA** en general, es explorar evidencia descriptiva y *configuracionalmente*, con miras hacia la diferentes formas en que las condiciones *causalmente relevantes*, pueden combinarse para producir un determinado resultado. Este capítulo proporciona a los investigadores interesados en la causalidad compleja, una variedad de estrategias y herramientas para descubrirlo y analizarlo, al mismo tiempo que acerca a los investigadores a sus casos y su evidencia.

## Resumen fsQCA para cómo hacer tablas de verdad con datos difusos (*fuzzy-sets*)

A continuación, lo listamos en la **Tabla 4.18.**

**Tabla 4.18. Resumen fsQCA para cómo hacer tablas de verdad con conjunto de datos difusos (*fuzzy-sets*)**

| Item | Etapa | Descripción |
|---|---|---|
| 1 | **Creando el conjunto de datos con puntuaciones difusas** | Cree un conjunto de datos con puntuaciones de pertenencia de conjunto difuso. (Conjuntos de datos nítidos pueden incluirse entre las *condiciones causales*.) Los conjuntos de *datos difusos* deben ser cuidadosamente etiquetados y definidos (por ejemplo, grado de pertenencia en el conjunto de empresas innovadoras cluster con niveles de de calidad de votación, persistentemente bajos). Preste mucha atención a:<br>**a.** La calibración de las puntuaciones de pertenencia difusos, especialmente con respeto a los *tres refrencias cualitativos*: *pertenencia completa, no pertenencia completa, y el punto de cruce*.<br>**b.** Los conjuntos difusos a menudo son unimodales a 1.0 o 0.0, o bimodal en **0.0** y **1.0**.<br>**c.** En general, la calibración requiere una base sólida de conocimiento del marco teórico y empírico sustantivo, así como comprensión profunda de los casos.<br>**d.** Los procedimientos descritos en este capítulo, funcionan mejor cuando la puntuación de pertenencia **0.5** y los más cercanos a **0.5**, se usan con moderación, (preferiblemente para todos), cuando se codifican las *condiciones causales*. |
| 2 | **Ingreso de datos directamente a programa (fsQCA)** | Ingrese los *datos difusos* directamente dentro del porgrama **fsQCA** (o similar), que pueda<br>**a.** Guardar archivos de datos en un formato compatible con **fsQCA** (por ejemplo, **Excel** usando archivos delimitados por comas, o **SPSS**, usando archivos delimitados por tabulaciones; los nombres de variables simples sin espacios incrustados o puntuación constituyen la primera fila del archivo de datos).<br>**b.** El conjunto de datos debe incluir tanto el resultado como las *condiciones causales relevantes*.<br>**c.** Abrir el archivo de datos usando **fsQCA** versión 2.0 o posterior. (Haga clic en Ayuda (*Help*) en la pantalla de inicio para identificar la versión y fecha de **fsQCA**; la versión más actualizada puede ser descargado de www.fsqca.com). |

| | | |
|---|---|---|
| **3** | **Selección preliminar de la lista de condiciones causales** | En general:<br>**a.** El número de *condiciones causales* debe ser modesto, en el rango de *tres a ocho*.<br>**b.** A menudo, las *condiciones causales* se pueden combinar de alguna manera para crear *macrovariables* utilizando los procedimientos descritos en **Ragin (2000)**.<br>**c.** Estas *macrovariables* se pueden usar en lugar de su componentes para reducir la dimensionalidad del espacio vectorial. Por ejemplo, una sola *macrovariable* podría usarse para reemplazar tres *condiciones causales* sustituibles unidas por lógica **OR**, que dicta el uso de su puntuación máximo de pertenencia.<br>**d.** Para realizarlo: en la ventana hoja de cálculo de datos de **fsQCA**, haga clic en Variables (*Variables*), posteriormente Calcular (*Compute*), y posteriormente use la función *fuzzyor* para crear este tipo de macrovariable). |
| **4** | **Creación de la tabla de verdad** | Cree una *tabla de verdad* especificando:<br>**a.** El resultado y las condiciones de causalidad.condiciones.<br>**b.** En **fsQCA**, se accede a esta función haciendo clic en Analizar (*Analyze*); conjuntos difusos (*Fuzzy Sets*) y algoritmo de *tabla de verdad* (*Truth Table Algorithm*).<br>**c.** La *tabla de verdad* resultante deberá tener $2^k$ filas, que reflejan las diferentes *esquinas del espacio vectorial*. (Los 1s y 0s para las condiciones causales, identifican las diferentes *esquinas de el espacio vectorial*.)<br>**d.** Para cada fila, el programa informa el número de casos con más de **0.5** de pertenencia en la esquina del espacio vectorial (en la columna etiquetada como *number*).<br>**e.** Dos columnas a la derecha del número (*number*) está la columna etiquetada como *consistencia* (*consistency*), que evalúa el grado en que la pertenencia de cada esquina, es un subconjunto de la pertenencia del resultado, para todos los casos. |

| | | |
|---|---|---|
| **5** | **Selección de la frecuencia de umbral** | El investigador deberá:<br>**a.** Seleccionar un umbral de frecuencia para aplicar al datos listados en la columna de números (*number column*).<br>**b.** Cuando el número total de casos incluidos en un estudio, es relativamente pequeño, el umbral de frecuencia debe ser 1 o 2.<br>**c.** Cuando el número N total es grande, sin embargo, un umbral más sustancial debe ser seleccionado.<br>**d.** Es muy importante inspeccionar la distribución de los casos al decidir un umbral de frecuencia. Esto puede ser logrado simplemente haciendo clic en cualquier caso en la columna de números (*number column*) y posteriormente haciendo clic en el menú Ordenar (*Sort*) y posteriormente seleccionar Descendiente (*Descending*).<br>**e.** La lista resultante de la cantidad de casos con más de **0.5** de pertenencia en cada esquina del espacio vectorial, proporcionará una instantánea de la distribución y también puede revelar discontinuidades o brechas importantes.<br>**f.** Después de seleccionar un umbral, elimine todas las filas que no la cumplen.<br>**g.** Esto se puede lograr (para tablas que se han ordenado según el la columna *number*) haciendo clic en el primer caso que cae por debajo del umbral (en la columna de *number*), haciendo clic en el menú Editar (Edit) y posteriormente clic en Eliminar fila actual hasta la ultima (*Delete current row to last*).<br>**h.** La verdad la tabla ahora enumerará, solo las *esquinas del espacio vectorial* que cumplen con el umbral de frecuencia. |

| | | |
|---|---|---|
| **6** | **Selección de umbral de consistencia** | La selección de un umbral de *consistencia* se utiliza para para distinguir combinaciones causales que son subconjuntos del resultado de aquellos que no lo son. Esta determinación se realiza utilizando:<br>**a.** La medida de la *consistencia* del marco teórico reportada en la columna de *consistencia* (*consistency column*).<br>**b.** En general, valores inferiores a **0.75** en dicha columna, indican una *inconsistencia sustancial*.<br>**c.** Es siempre útil, ordenar las puntuaciones de *consistencia* en orden descendente para que que sea posible evaluar su distribución.<br>**d.** Esto debe hacerse después de que se hayan eliminado las filas que caen por debajo del umbral de frecuencia de la tabla (paso 5).<br>**e.** Haga clic en cualquier valor en la columna de *consistencia* (*consistency column*); haga clic en el menú Ordenar (*Sort*); y posteriormente haga clic en Descendente (*Descending*).<br>**f.** Identifique cualquier brecha, en el rango superior de la *consistencia* que podría ser útil para establecer un umbral, teniendo en cuenta que siempre es posible examinar varios umbrales diferentes y evaluar las consecuencias de bajar y elevar el corte de *consistencia*.<br>**g.** A menudo, es útil presentar dos análisis, uno con un umbral de *consistencia* relativamente permisivo (por ejemplo, alrededor de 0.8) y otro con un umbral de *consistencia* más restrictivo (por ejemplo, alrededor de **0.9**). |
| **7** | **Prueba preliminar** | Para lograrlo:<br>**a.** Ingrese manualmente 1s y 0s dentro de la columna de resultado vacía, (o que se encuentra etiqueta con el nombre del resultado) y listada a la izquierda de la columna de *consistencia* (*consistency column*).<br>**b.** Usando el valor umbral seleccionado en el paso anterior, ingrese un valor de 1 cuando el valor de *consistencia* cumple o excede el umbral de *consistencia* y 0 en caso contrario.<br>**c.** Si la tabla de la verdad, en la hoja de cálculo, tiene más de veinte filas, codifique la columna de resultados usando la función Eliminar y Codificar (*Delete and Code*) en el menú Editar (*Edit*). |
| **8** | **Análisis Estándar** | Para realizarlo:<br>**a.** Haga clic en el botón Análisis Estándar (*Standard Analyses*) en la parte inferior de la pantalla.<br>**b.** Se producen tres soluciones: la compleja *(complex)*, la *parsimoniosa (parsimonious)* y la intermedia *(intermediate)*.<br>**c.** Estas diferentes soluciones se explican en los capítulos 8 y 9. La solución *intermedia (intermediate)* se basa en información sobre las condiciones de causalidad que el usuario ingresa, con base a su conocimiento sustantivo. |

Fuente: Ragin (2008) con adaptación propia.

## Casos fsQCA

A continuación, se presentarán una serie de ejercicios con el software **fsQCA 3.0**, para que lo conozca y comprenda los conceptos hasta aquí vertidos.

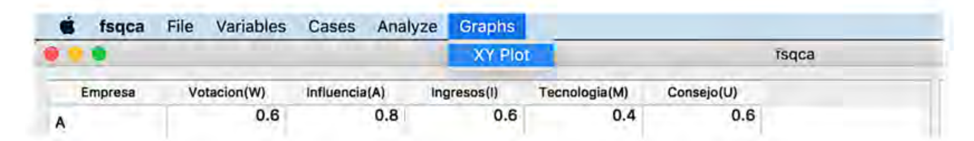
### Gráfica con datos en negación. Caso 1

Del archivo de trabajo: **empresas innovadoras cluster.cvs (Tabla 4.5)** realice gráfico **W** vs. **(~ U)** explique y documente como **Gráfica 4.1** Para lograrlo:

1. Abra el archivo **empresas innovadoras cluster.cvs**, oprimiendo**: Open → File**

| Empresa | Votacion(W) | Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) |
|---|---|---|---|---|---|
| A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 |
| C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 |
| D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 |
| E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 |
| F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 |
| G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 |
| H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 |
| I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 |
| J | 0 | 0.8 | 0.4 | 0.8 | 1 |
| K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 |
| L | 1 | 1 | 0.8 | 0.4 | 0.2 |

2. Para graficar, oprima: **Graphs→XY Plot**

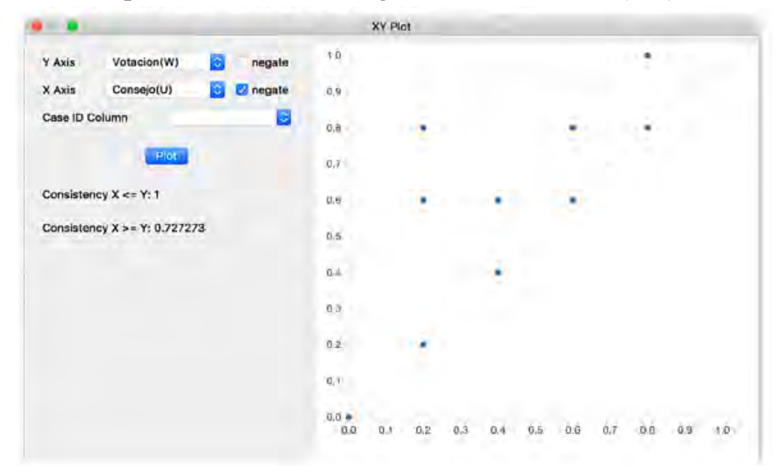


**3.** Coloque en las casillas **Y Axis** = *Votacion,* **W**; **X Axis** = *Consejo*, **U** y marcar **negate;** oprimir **Plot**




**4.** A continuación se desplegará el gráfico a marcar como, y se procede a su explicación:

**Gráfica 4.1. Conjunto de datos difusos (*fuzzy-sets*) que muestra un subconjunto de datos W vs. (~ U)**




Este gráfico difuso es muy importante, dadas sus implicaciones sustantivas. **Schneider y Grofman (2006)** advierten, *no se trata de una ecuación de regresión aunque lo parezca*. El **XY Plot** permite comprender las relaciones entre las *condiciones y el resultado* (o entre condiciones entre sí) y, en particular, la *cobertura* y *consistencia* de estas relaciones. Como afirma **Ragin (2008, p. 60)**, el **XY Plot** muestra de forma intuitiva las relaciones entre conjuntos, pueden ser:

- Cuando todos o casi todos los casos caen por encima de la diagonal esto indica una relación de *suficiencia.*
- Mientras que si todos o casi todos los casos caen por debajo de la diagonal esto indica una relación de *necesidad*.

La característica de la forma gráfica triangular superior izquierda, indica que el conjunto trazado en el eje horizontal (conjunto de empresas innovadoras del clúster con fortaleza resultados consejo de administración, **~ U**) es un subconjunto del conjunto trazado en el eje vertical (conjunto de empresas innovadoras del clúster con calidad de votación por liderazgo tecnológico, **W**). Los puntos en la región superior izquierda de la gráfica, no son errores como se considerarían en un análisis de regresión lineal. Más bien, estos puntos tienen una fuerte pertenencia en el resultado debido a la operación de otras *condiciones causales* u otras combinaciones de *condiciones causales*. El triángulo inferior vacante en esta gráfica de conjuntos de *datos difusos,* corresponde para vaciar la celda **C4** de la **Tabla 4.8**, que utiliza el conjunto de datos nítidos. Así como los casos en la celda **C4**, violarían los subconjuntos de datos nítidos, los casos en la parte inferior derecha el triángulo de la **Gráfica 4.1,** violarían la relación de subconjunto difuso.

La **Tabla 4.5** ilustran la relación de un subconjunto de *datos difusos*, utilizando una sola *condición causal*. Tenga en cuenta, sin embargo, que esta misma evaluación podría haberse realizado utilizando el grado de pertenencia de una combinación de *condiciones causales*. Como se señaló anteriormente, para calcular el grado de pertenencia de un caso en una combinación de condiciones, es necesario simplemente usar el puntaje de pertenencia más bajo (mínimo) entre las *condiciones causales,* que se desprende de la aplicación de la lógica y la operación de álgebra difusa. El grado de pertenencia en una combinación causal, se puede utilizar para evaluar la relación del subconjunto de *datos difusos,* mediante la comparación de puntuaciones en la combinación causal (eje horizontal) con puntajes de pertenencia en el resultado (eje vertical). Este examen establece, si el grado de pertenencia en una combinación de condiciones de causalidad es un subconjunto de *datos difusos,* del grado de pertenencia en el resultado,

tiene un patrón de resultados consistente con un argumento de causalidad suficiente **(Ragin, 2000)**. Un diagrama triangular superior izquierdo, con grado de pertenencia en la combinación causal en el eje horizontal y el grado de pertenencia en el resultado en el eje vertical, señala la difusa relación de subconjunto.

Recordemos varios de los puntos principales sobre las relaciones establecidas en la investigación social, presentado en el capítulo 1, a saber:

1. Establecen relaciones a menudo implican conexiones causales u otras integrales entre fenómenos sociales,
2. Son fundamentalmente asimétricas, y
3. Pueden ser fuertes a pesar de tener correlaciones relativamente débiles.

Considere la declaración: entre las empresas innovadoras del clúster, aquellas que carecen de fortaleza de resultados del *consejo de administración* (**U**) tienen débil *calidad de votación por liderazgo tecnológico* **W.** La declaración plantea la hipótesis de un enlace explícito entre **U** y **W.** Como muchas de estas declaraciones, primero se presenta al subconjunto (*calidad de votación por liderazgo tecnológico,* **W**) y afirma, en esencia, que el conjunto de empresas innovadoras en el clúster que carecen de fortaleza de resultados del *consejo de administración* **U,** constituyen un subconjunto de las empresas innovadoras del clúster con calidad de *votación por liderazgo tecnológico,* **W.** La declaración es fundamentalmente asimétrica. No afirma que las empresas innovadoras del clúster con fortaleza de resultados *del consejo de administración* **U**, estén de alguna manera excluidos o que les impida tener una votación de *calidad por liderazgo tecnológico* **W**, y por lo tanto, deja abierta la posibilidad de que existan otros obstáculos para **W.** Tal evidencia, no cuestiona directamente la afirmación de que hay un camino a través de débiles consejos de administración.

Finalmente, desde un punto de vista del marco teórico de soporte, la evidencia presentada es *perfectamente* consistente con la afirmación teórica afirma: todos los casos que están en el triángulo superior izquierdo de la gráfica. Desde un punto de vista correlacional, sin embargo, la evidencia no es tan perfecta, dada la correlación de Pearson r = **0.766.**

## Principio de subconjunto y relación aritmética entre las puntuaciones de pertenencia en datos nítidos (*crisp-sets*)

Considere nuevamente el ejemplo de la **Tabla 4.5.**

**Tabla 4.5. Datos de conjunto de datos difusos (*fuzzy-sets*) para el caso de decisión de clúster de innovación**

| Empresas innovadoras Clúster | Columnas | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | Calidad de votación por liderazgo tecnológico | Nivel de influencia | Diferencia de ingresos por innovación | Grado de tecnología manufactura | Fortaleza resultados consejo de administración | (~ M) | (A) * ( ~M) | (A) + (~M) |
| | W | A | I | M | U | | | |
| **A** | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 | 0.6 | 0.6 | 0.8 |
| **B** | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 | 0.8 | 0.6 | 0.8 |
| **C** | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 | 0.8 | 0.6 | 0.8 |
| **D** | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 | 0.8 | 0.6 | 0.8 |
| **E** | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 | 0.6 | 0.6 | 0.6 |
| **F** | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 | 0.2 | 0.2 | 0.2 |
| **G** | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 | 0.8 | 0.4 | 0.8 |
| **H** | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 | 0.8 | 0.6 | 0.8 |
| **I** | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 | 0.4 | 0.4 | 0.6 |
| **J** | 0.0 | 0.8 | 0.4 | 0.8 | 1.0 | 0.2 | 0.2 | 0.8 |
| **K** | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 | 0.2 | 0.2 | 0.6 |
| **L** | 1.0 | 1.0 | 0.8 | 0.4 | 0.2 | 0.6 | 0.6 | 1.0 |

Fuente: Elaboración propia.

Donde lograr una diferencia en sus *ingresos por innovación* **I,** es una condición necesaria pero no suficiente para lograr pertenecer al conjunto de empresas innovadoras con *calidad de votación por liderazgo tecnológico*, **W.** Se deduce lógicamente que:

**1.** Si una condición es necesaria pero no suficiente para un resultado, las instancias del resultado constituirán un subconjunto de instancias

de la causa. Otra forma de entender la relación del subconjunto, es en términos de la relación aritmética entre los puntuaciones de pertenencia de los datos nítidos **(1s y 0s),** o sea,

2. Si las instancias del resultado, son un subconjunto de instancias de la causa, entonces el valor booleano del resultado **(1 vs. 0),** será menor o igual que el valor booleano de la causa.

## Principio de subconjunto y relación aritmética entre las puntuaciones de pertenencia en datos difusos (*fuzzy-sets*)

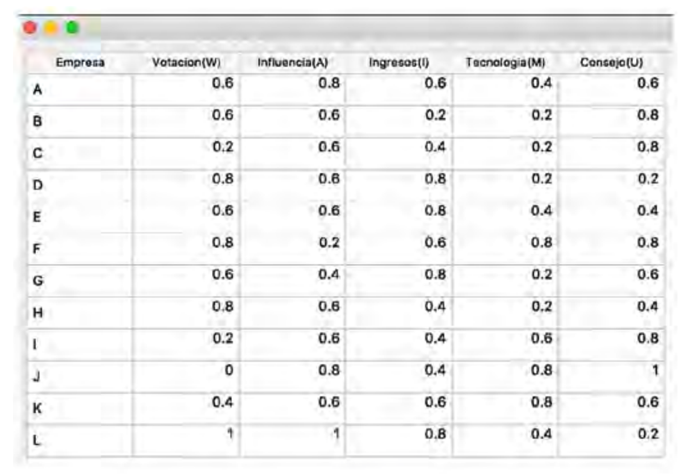
Con conjuntos de *datos difusos*, sería difícil seleccionar empresas innovadoras con *calidad de votación por liderazgo tecnológ*ico, **W** con el resultado (primer paso habitual, en el análisis de datos nítidos, de condiciones nece-sarias) porque las empresas del ejemplo, varían en su grado de pertenencia en *fortaleza de resultados del consejo de administración,* **U.** Del mismo modo, sería muy difícil evaluar, de acuerdo con los casos con respecto a la *condición causal* relevante (empresas innovadoras con *calidad de votación por liderazgo tecnológico,* **W**), debido a también varían en su pertenencia en este conjunto.

## Gráfica con datos sin negación. Caso 2

Afortunadamente, el principio del subconjunto y la relación aritmética entre las puntuaciones de pertenencia en conjuntos de datos nítidos, también son válidos para los difusos. En estos, el conjunto **A** es un subconjunto del conjunto **B**, si los puntuaciones de pertenencia de los casos en el conjunto **A** son menores o iguales, a los puntuaciones de pertenencia respectivos en el conjunto **B**. Del archivo **empresas innovadoras cluster. cvs** (**Tabla 4.5**) realice gráfico **W** vs. **U** y genere **Gráfica 4.2**. explicando su significado.

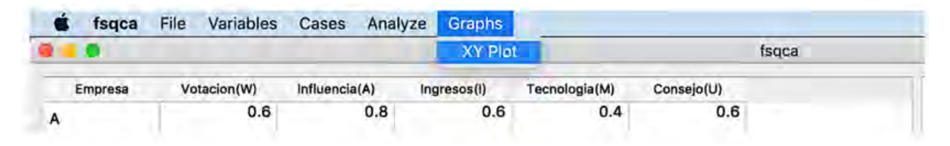
La **Gráfica 4.2** ilustra la relación de subconjunto difuso usando los puntajes de la **Tabla 4.5.** Para lograrlo:

**1.** Abra el archivo de trabajo: **empresas innovadoras cluster.cvs**, oprimiendo: **Open → File**



| Empresa | Votacion(W) | Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) |
|---|---|---|---|---|---|
| A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 |
| C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 |
| D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 |
| E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 |
| F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 |
| G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 |
| H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 |
| I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 |
| J | 0 | 0.8 | 0.4 | 0.8 | 1 |
| K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 |
| L | 1 | 1 | 0.8 | 0.4 | 0.2 |

**2.** Para graficar, oprima: **Graphs → XY Plot**

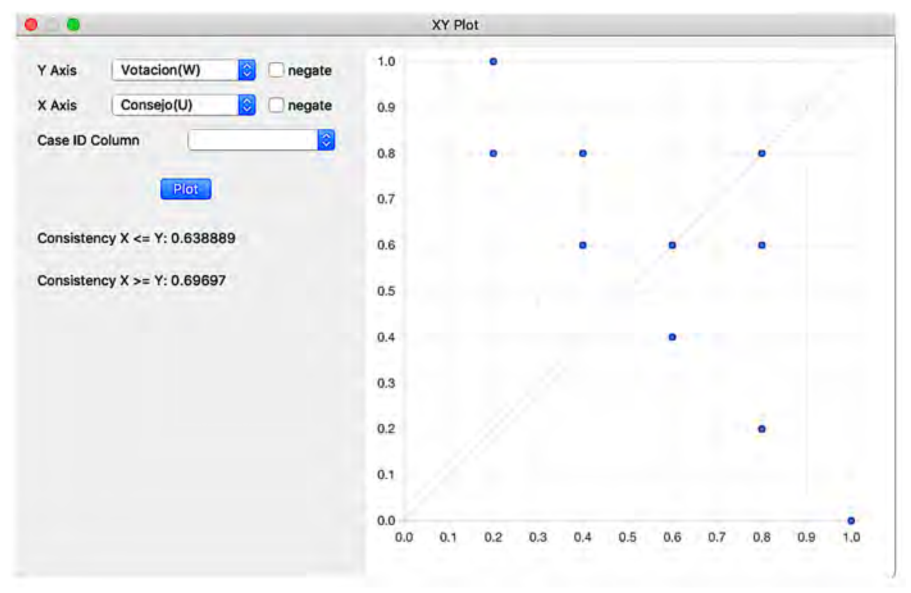



**3.** Coloque en las casillas **Y Axis**= *Votacion,* **W**; **X Axis**= *Consejo,* **U** y oprimir **Plot**



4. A continuación se delegará el gráfico a marcar como **Gráfica 4.2**, y se procede a su explicación.

**Gráfica 4.2. Gráfica de conjunto de datos difusos (*fuzzy-sets*) que muestra un subconjunto de dato W vs. U**




## Cuando la puntuación es necesaria

Cuando las puntuaciones de pertenencia del conjunto de datos difuso en el resultado, son menores o iguales a los de la pertenencia difusa en la causa, entonces es posible argumentar, que las instancias del resultado, son un subconjunto de instancias de la causa. La **Gráfica 4.2**, muestra esta relación aritmética en dos dimensiones. Cuando los investigadores encuentran este patrón, podrían citar esta evidencia, como respaldo para un argumento de una *necesidad causal.*

## Cuando la puntuación es suficiente

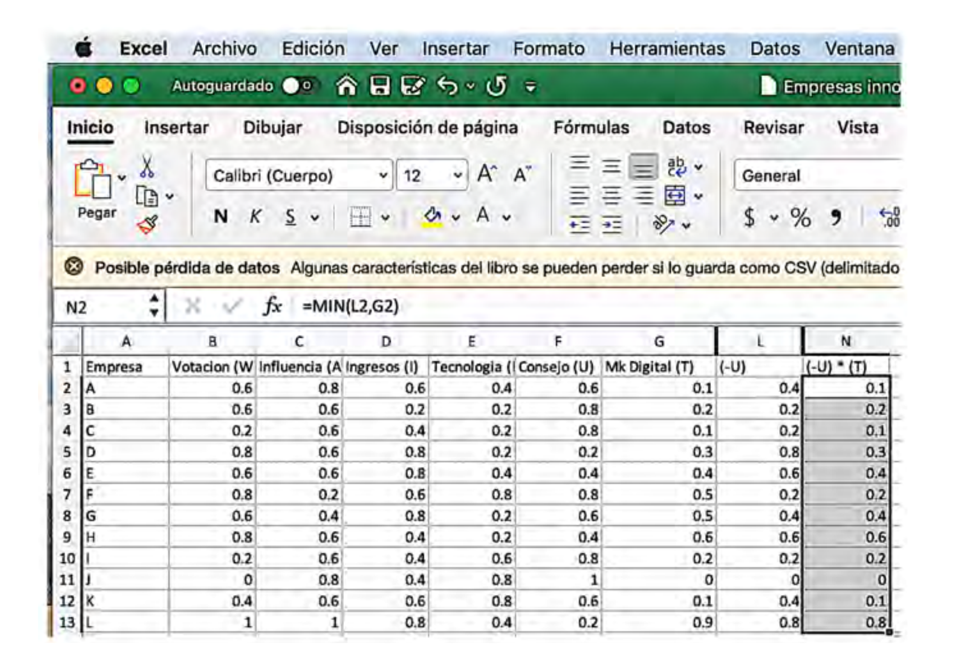
La evaluación de suficiencia, puede verse como una prueba de si los casos que muestran las **condiciones causales**, forman un subconjunto de los casos que muestran el resultado. Como se muestra arriba, otra forma de entender la relación del subconjunto, es en términos de la relación aritmética entre las puntuaciones de pertenencia.

## Gráfica considerando intersección de casos. Caso 3

Para argumentar que una causa o combinación causal, es *suficiente* para el resultado, las puntuaciones de pertenencia del conjunto de datos difuso, en la causa, deben ser menores o iguales a las puntuaciones de pertenencia del conjunto de datos difusa, en el resultado.

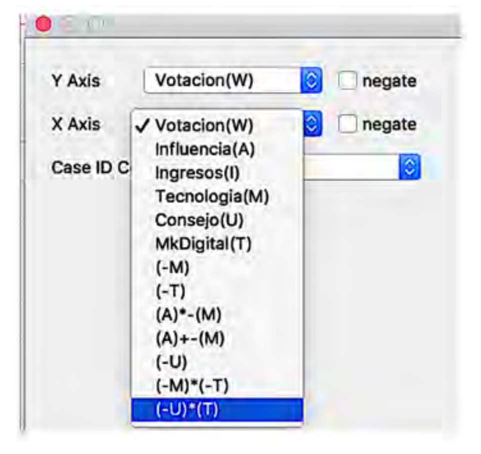
Así, del archivo de trabajo: **empresas innovadoras cluster.cvs** (**Tabla 4.5**) realice gráfico **W** vs. **(~U) * (T),** y genere **Gráfica 4.3**. explicando su significado, la cual debe ser el resultado de la relación aritmética entre las combinaciones causales suficientes de, empresas innovadoras con *calidad de votación por liderazgo tecnológico* **W** vs. la no fortaleza de resultados del consejo de administración, **(~U) AND** el *uso de la innovación por mercadotecnia digital,* **T.** Para lograr la gráfica:

1. Del archivo de trabajo: **empresas innovadoras cluster.cvs,** realizar la intersección **(~U * T),** esto se realiza aplicando la función **MIN** entre las columnas respectivas en **Excel**.

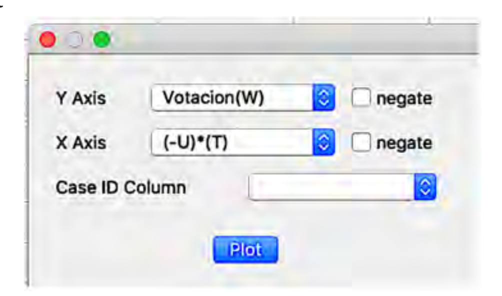


| | A | B | C | D | E | F | G | L | N |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Empresa | Votacion (W | Influencia (A | Ingresos (I) | Tecnologia (I | Consejo (U) | Mk Digital (T) | (-U) | (-U) * (T) |
| 2 | A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 | 0.1 | 0.4 | 0.1 |
| 3 | B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 | 0.2 | 0.2 | 0.2 |
| 4 | C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 | 0.1 | 0.2 | 0.1 |
| 5 | D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 | 0.3 | 0.8 | 0.3 |
| 6 | E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 | 0.4 | 0.6 | 0.4 |
| 7 | F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 | 0.5 | 0.2 | 0.2 |
| 8 | G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 | 0.5 | 0.4 | 0.4 |
| 9 | H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 | 0.6 | 0.6 | 0.6 |
| 10 | I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 | 0.2 | 0.2 | 0.2 |
| 11 | J | 0 | 0.8 | 0.4 | 0.8 | 1 | 0 | 0 | 0 |
| 12 | K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 | 0.1 | 0.4 | 0.1 |
| 13 | L | 1 | 1 | 0.8 | 0.4 | 0.2 | 0.9 | 0.8 | 0.8 |

2. Abrir archivo desde **fsQCA,** repitiendo el **pasos 1** del **ejercicio 1 y 2** anteriores. Así, para graficar, oprina en la casilla **X Axis**= **W y Y Axis= (~U) * (T)**

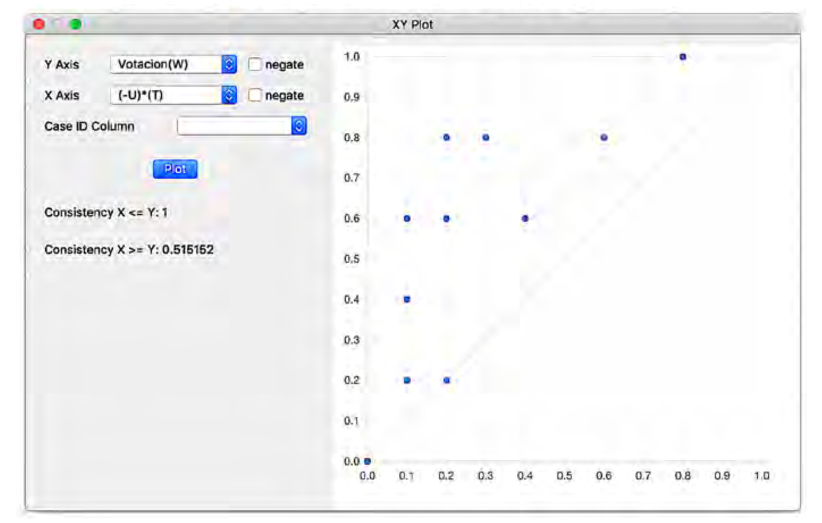


**3.** Oprimir **Plot**




**4.** A continuación se delegará el gráfico a marcar como **Gráfica 4.3**, y se procede a su explicación.

**Gráfica 4.3. Gráfica de conjunto de datos difusos (*fuzzy-sets*) que muestra un subconjunto de dato (W) vs. (~U) * (T)**

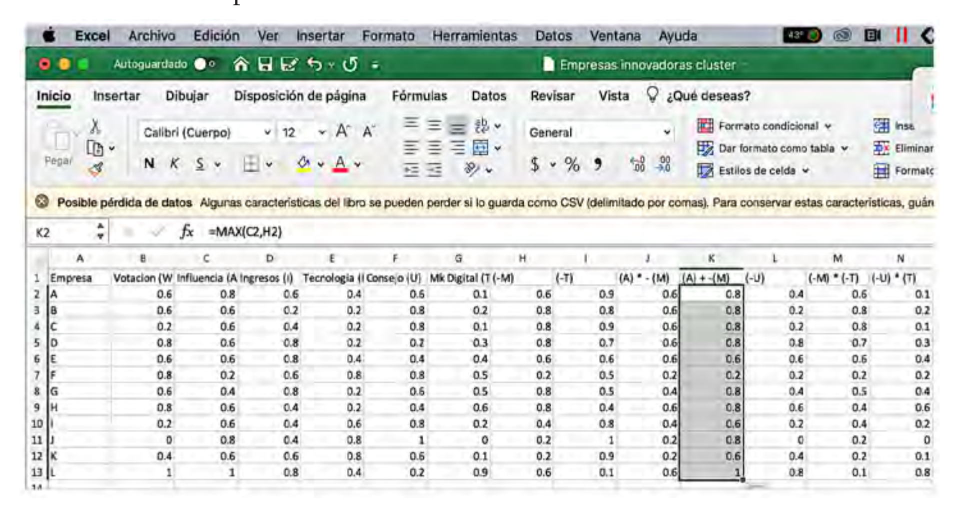



**5.** El diagrama triangular superior que se muestra en la **Gráfica 4.3** es un reflejo directo, del hecho de que los puntajes de membresía en la empresas innovadoras con calidad de *votación por liderazgo tecnológico*, **W** del conjunto de datos difuso, son menores o iguales que los puntajes de pertenencia en la *no fortaleza de resultados del consejo de administración*,**( ~U) AND** el uso de la *innovación por mercadotecnia digital*, **T.**

## Gráfica considerando unión de casos. Caso 4

Para argumentar que una causa o combinación causal, es *necesaria* para el resultado, las puntuaciones de pertenencia del conjunto de datos difuso en la causa, deben ser mayores o iguales a las puntuaciones de pertenencia del conjunto de datos difusa, en el resultado. Del archivo de trabajo: **empresas innovadoras cluster.xls** (**Tabla 4.5**) realice gráfico **W** vs. **( A + ~M),** y genere **Gráfica 4.4**. explicando su significado, la cual es resultado de la relación aritmética entre las combinaciones causales suficientes de empresas innovadoras con calidad de votación por liderazgo tecnológico, **W** vs. el grado de pertenencia, en el conjunto de empresas innovadoras clúster por su nivel de influencia **A OR** el grado bajo de pertenencia, en el conjunto empresas innovadoras clúster con empleo de un alto grado tecnológico de manufactura **(~M).** Para lograr la gráfica,realice:

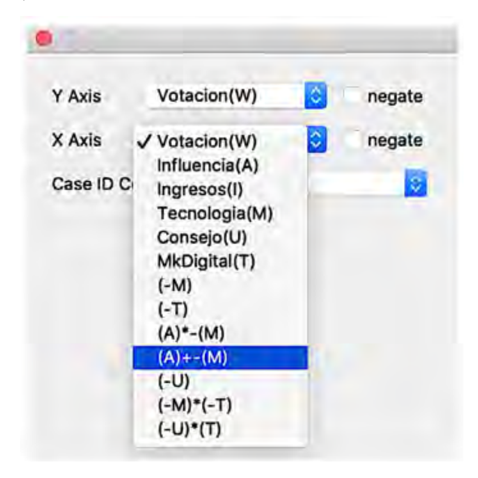
1. Del archivo de trabajo: **empresas innovadoras cluster.xls,** realizar la unión de **A + (~M),** esto se realiza aplicando la función **MAX** entre las columnas respectivas en **Excel**.



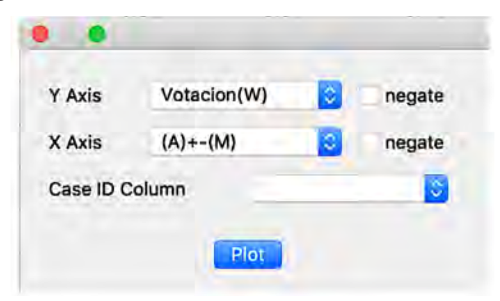
K2 =MAX(C2,H2)

| Empresa | Votacion (W | Influencia (A | Ingresos (I) | Tecnologia (I | Consejo (U) | Mk Digital (T | (-M) | (-T) | (A) * - (M) | (A) + -(M) | (-U) | (-M) * (-T) | (-U) * (T) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 | 0.1 | 0.6 | 0.9 | 0.6 | 0.8 | 0.4 | 0.6 | 0.1 |
| B | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 | 0.2 | 0.8 | 0.8 | 0.6 | 0.8 | 0.2 | 0.8 | 0.2 |
| C | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 | 0.1 | 0.8 | 0.9 | 0.6 | 0.8 | 0.2 | 0.8 | 0.1 |
| D | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 | 0.3 | 0.8 | 0.7 | 0.6 | 0.8 | 0.8 | 0.7 | 0.3 |
| E | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 | 0.4 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.4 |
| F | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 | 0.5 | 0.2 | 0.5 | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| G | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 | 0.5 | 0.8 | 0.5 | 0.4 | 0.8 | 0.4 | 0.5 | 0.4 |
| H | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 | 0.6 | 0.8 | 0.4 | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| I | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 | 0.2 | 0.4 | 0.8 | 0.4 | 0.6 | 0.2 | 0.4 | 0.2 |
| J | 0 | 0.8 | 0.4 | 0.8 | 1 | 0 | 0.2 | 1 | 0.2 | 0.8 | 0 | 0.2 | 0 |
| K | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 | 0.1 | 0.2 | 0.9 | 0.2 | 0.6 | 0.4 | 0.2 | 0.1 |
| L | 1 | 1 | 0.8 | 0.4 | 0.2 | 0.9 | 0.6 | 0.1 | 0.6 | 1 | 0.8 | 0.1 | 0.8 |

2. Abrir archivo desde **fsQCA,** repitiendo el paso 1, de los ejercicios 1 y 2 anteriores. Así, para graficar, oprina en la casilla **X Axis**= **W y Y Axis= A + (~M)**

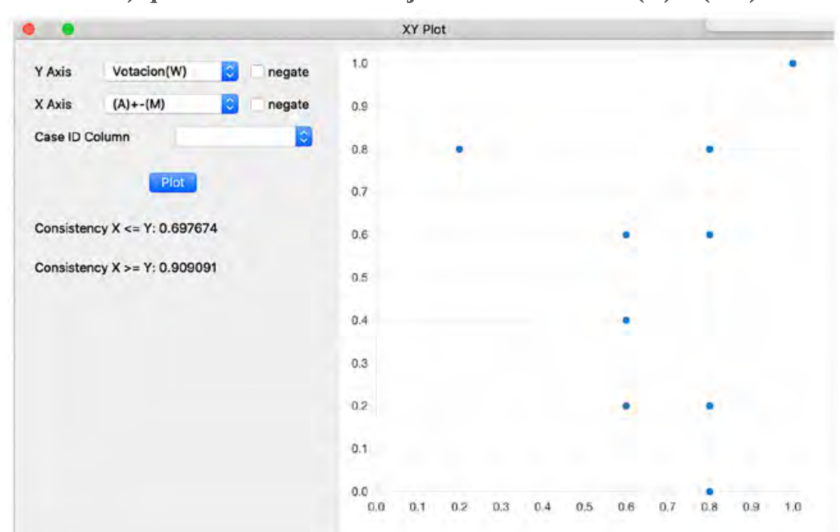



3. Oprimir **Plot**




4. A continuación se delegará el gráfico a marcar como **Gráfica 4.4**, y se procede a su explicación.

**Gráfica 4.4. Gráfica de conjunto de datos difusos (*fuzzy-sets*) que muestra un subconjunto de dato W vs. (A) + (~M)**




Donde el resultado **W,** es un subconjunto de la *condición causal* **A + (~M)** ; por lo tanto, todos los valores de **W,** son menores o iguales que sus valores **A + (~M)** correspondientes. Tenga en cuenta, que los casos en la esquina inferior derecha de la gráfica no contradicen el es necesario, para estos son casos, que tienen baja pertenencia en el resultado porque carecen de alguna otra *condición causal* no especificada. Después de todo, lo *condición causal* en este ejemplo solo *es necesaria, pero no suficiente.*

## Uso del algoritmo de tablas de verdad de datos difusos(*fuzzy-sets*)

El algoritmo de la *tabla de verdad* del conjunto de *datos difusos*, puede conceptualizarse como un puente con tres pilares:

**1.** El primer pilar, es la correspondencia directa que existe, entre las filas de una *tabla de verdad* del conjunto de datos nítidos y las coordenadas

de las *esquinas del espacio vectorial*, definidas por *condiciones causales* establecidas difusas **(Ragin, 2000)**.

2. El segundo pilar, es la evaluación de la distribución de casos entre diferentes combinaciones lógicamente posibles, de *condiciones causales* (o sectores del espacio vectorial). Algunos sectores del espacio vectorial, pueden tener muchos casos con pertenencia fuerte, mientras que otros sectores pueden tener casos con pertenencia débil.
3. El tercer pilar es la evaluación de la *consistencia* de la evidencia para cada combinación causal, con el argumento de que es un subconjunto difuso del resultado. El algoritmo de la *tabla de verdad*, implica establecer estos tres pilares para construir, una *tabla de verdad* de datos nítida, donde el análisis procede de manera similar, al algoritmo del conjunto de datos nítidos.

## Análisis de tabla de verdad de datos difusos (*fuzzy-sets*). Caso 5

Esta sección explicará los pasos necesarios para registrar los resultados del análisis de conjuntos de *datos difusos*, en una *tabla de verdad* del conjunto de datos nítidos y luego analizar esta tabla. Los conjuntos de *datos difusos*, pueden importarse de otros programas o crearse en **fsQCA,** como se describe en el capítulo 2. Con el uso de la **Tabla 4.5** de empresas con *voto de calidad por liderazgo tecnológico,* **W** en un clúster de innovación y que corresponde al de **Empresas innovadoras cluster.xls**. mostrada abajo, despliegue y analice la *tabla de verdad*.

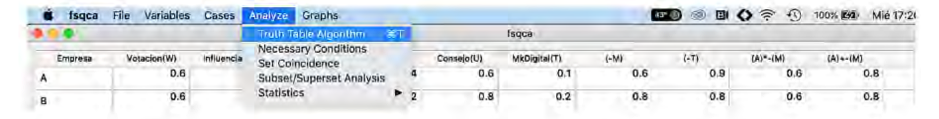
| | Empresa | Votacion(W) | Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) |
|---|---|---|---|---|---|---|
| A | | 0.6 | 0.8 | 0.6 | 0.4 | 0.6 |
| B | | 0.6 | 0.6 | 0.2 | 0.2 | 0.8 |
| C | | 0.2 | 0.6 | 0.4 | 0.2 | 0.8 |
| D | | 0.8 | 0.6 | 0.8 | 0.2 | 0.2 |
| E | | 0.6 | 0.6 | 0.8 | 0.4 | 0.4 |
| F | | 0.8 | 0.2 | 0.6 | 0.8 | 0.8 |
| G | | 0.6 | 0.4 | 0.8 | 0.2 | 0.6 |
| H | | 0.8 | 0.6 | 0.4 | 0.2 | 0.4 |
| I | | 0.2 | 0.6 | 0.4 | 0.6 | 0.8 |
| J | | 0 | 0.8 | 0.4 | 0.8 | 1 |
| K | | 0.4 | 0.6 | 0.6 | 0.8 | 0.6 |
| L | | 1 | 1 | 0.8 | 0.4 | 0.2 |

El algoritmo de la *tabla de verdad*, incorpora un procedimiento analítico de dos etapas:

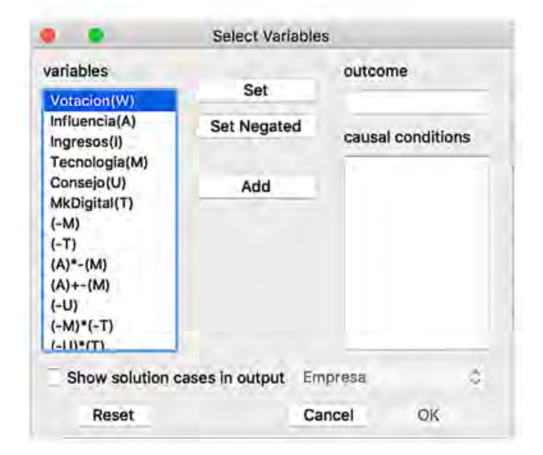
**a.** El primer paso, consiste en crear una *tabla de verdad* a partir del conjunto de *datos difusos*, que incluye, especificar el resultado para cada configuración y determinar qué *configuraciones* incluir en el análisis.
**b.** El segundo paso, consiste en especificar las *condiciones causales* y los resultados para minimizar. Estos pasos deben realizarse en conjunto y ambos deben realizarse para cada análisis por separado.

Para lograrlo:

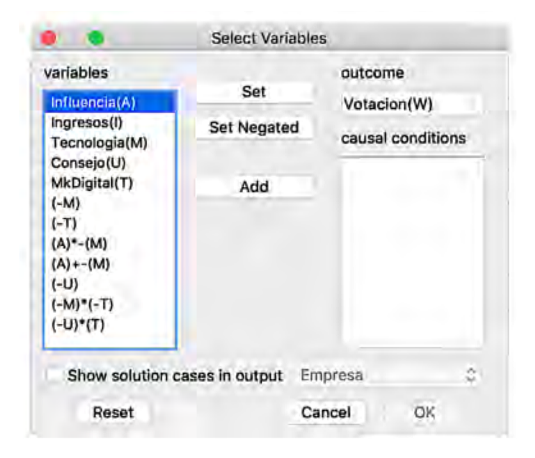
1. Abra el archivo referido, de acuerdo con el **paso 1** del **ejercicio 1 y 2** anteriores.
2. Oprima: **Analyze → Truth Algorithm**



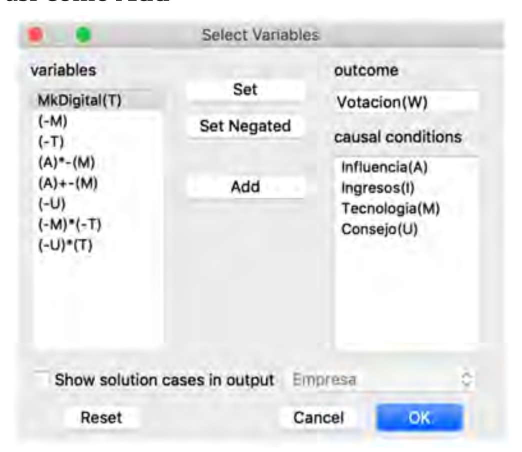
**3.** Del que aparecerá la siguiente ventana




**4.** Identifique, resalte y transfiera en la casilla correspondiente, las variables que desee usar como **Outcome** y marque **Set**



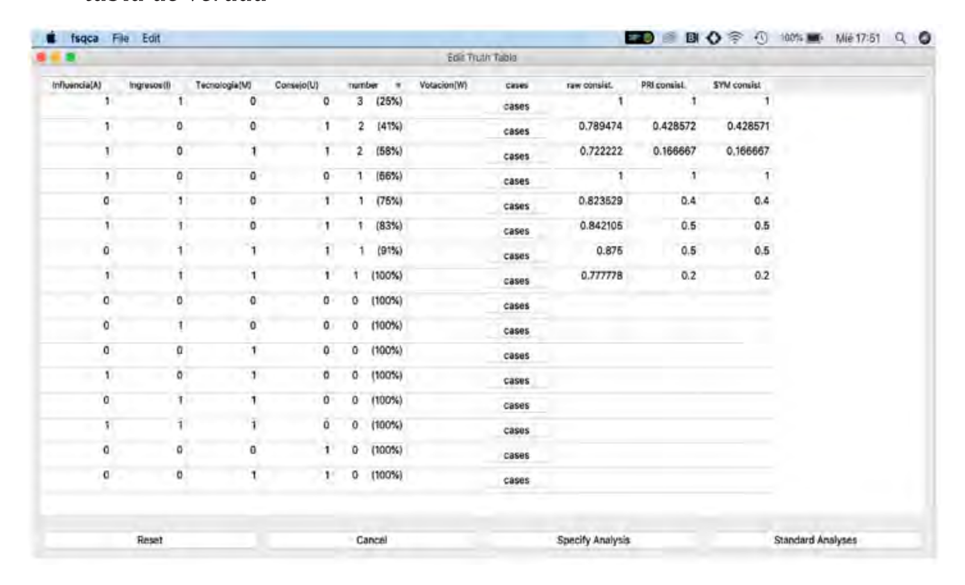
5. Seleccione las *condiciones causales* y oprima sobre el campo **Causal Conditions** así como **Add**




6. Marque la casilla junto a **Show solution cases in output** y elija la variable en la cual ordenó sus casos (aquí, **Empresa**)



**7.** Oprima el botón **OK** y aparecerá la siguiente ventana que contiene la *tabla de verdad*




| Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 3 (25%) | | cases | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 | 2 (41%) | | cases | 0.789474 | 0.428572 | 0.428571 |
| 1 | 0 | 1 | 1 | 2 (58%) | | cases | 0.722222 | 0.166667 | 0.166667 |
| 1 | 0 | 0 | 0 | 1 (66%) | | cases | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 (75%) | | cases | 0.823529 | 0.4 | 0.4 |
| 1 | 1 | 0 | 1 | 1 (83%) | | cases | 0.842105 | 0.5 | 0.5 |
| 0 | 1 | 1 | 1 | 1 (91%) | | cases | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 1 | 1 | 1 (100%) | | cases | 0.777778 | 0.2 | 0.2 |
| 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 0 | 0 (100%) | | cases | | | |
| 1 | 0 | 1 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 1 | 0 | 0 (100%) | | cases | | | |
| 1 | 1 | 1 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 0 | 1 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 1 | 0 (100%) | | cases | | | |



**8.** Como se aprecia, la *tabla de verdad* tiene ***2***$^{k}$ filas (donde ***k*** representa el número de *condiciones causales* = **4**), reflejando todas las combinaciones posibles de *condiciones causales*. Los **(1)** y **(0)**, indican las diferentes *esquinas del espacio vectorial* definidas por las *condiciones causales*, del conjunto de datos difuso. Para cada fila, se crea un valor para cada una de las siguientes variables (Ver **Tabla 3.27**).

**Tabla 3.27. Descripción de valores que reporta fsQCA en su versión conjunto de datos nítidos (*crisp-sets*)**

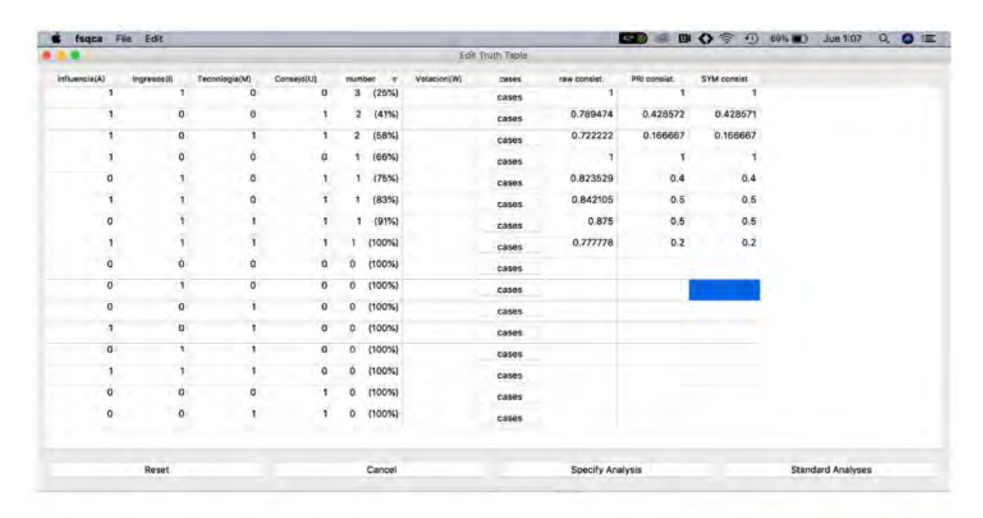
| Valor | Descripción |
|---|---|
| **number** | Es el número de casos, que muestran la combinación de condiciones, es decir, el número de casos existentes para cada configuración de condiciones. |
| **raw consist.** | Es la proporción de casos, en cada fila de la *tabla de verdad*, que muestran el resultado. Para comprender cuántos casos con el *resultado de interés* están cubiertos por cada configuración de condiciones, se utiliza este término como *índice de consistencia bruta*. Este índice indica qué porcentaje de casos dentro de una configuración de condiciones muestra el resultado de interés entre el número total de casos en dicha configuración. La *consistencia bruta* es importante ya que antes de realizar la minimización booleana debemos indicar cuáles de las configuraciones serán incluidas en la operación, y en particular debemos seleccionar el número de casos que las configuraciones deben tener para ser incluidos. La opción por default es (**1**) y los niveles de consistencia mínimos de las configuraciones que se incluirán. El nivel de consistencia mínimo aceptado rondaría **0.8** o incluso **0.85**, aunque niveles más altos de consistencia conducen a resultados más robustos, debiendo siempre ser explicitada y justificada la elección hecha a este respecto (véase para una discusión, **Ragin, 2004**). En particular, en **csQCA** es preferible incluir solo aquellas configuraciones con una consistencia de (**1**) pues niveles más bajos indican la existencia de una *contradicción*. En análisis de conjuntos difusos la interpretación del índice de consistencia bruta es más compleja, ya que en este tipo de análisis los casos tienen una inclusión difusa en cada una de las condiciones de la configuración **(Ragin, 2009)**. En el capítulo sobre, **fsQCA** se explica cómo se calcula e interpretan las relaciones de *suficiencia*. En caso de duda sobre el valor de *consistencia*, tanto para **csQCA** como para **fsQCA**, se habilitan dos parámetros complementarios: **PRI** y **SYM,** analizados en detalle por **Schneider y Wagernann (2012, pp. 242-243)**. |
| **PRI consist.** | Llamado *Propotional Reduction in Inconsistency*, es la medida alternativa de *consistencia* (desarrollada para *datos difusos*), basada en una reducción cuasi proporcional en el cálculo de errores. En análisis de datos nítidos, esto será igual a *raw-consist.* El **PRI**, que no puede entenderse como sustituto del valor de *consistencia*, indica en qué medida una determinada configuración forma parte de **Y** y no de (**-Y**) o, en terminología de conjuntos, en qué medida **X** es un subconjunto de **Y** y no de (**-Y**) (**Schneider y Wagemann, 2012, p. 242**). Siguiendo a los autores, se busca que las configuraciones sean consistentes con la idea de subconjuntos; así, un valor **PRI** bajo estará indicando que los valores de consistencia de (**X**) para (**Y**) e (**-Y**) están muy cerca, por lo que **X** fomará parte de ambos conjuntos. De lo contrario, se buscan valores altos en el parámetro **PRI**. |
| **SYM consist.** | Es la medida alternativa de *consistencia* para *datos difusos* basada en una versión simétrica de **PRI.** |

Fuente: **fsQCA 3.0** con adaptación propia.

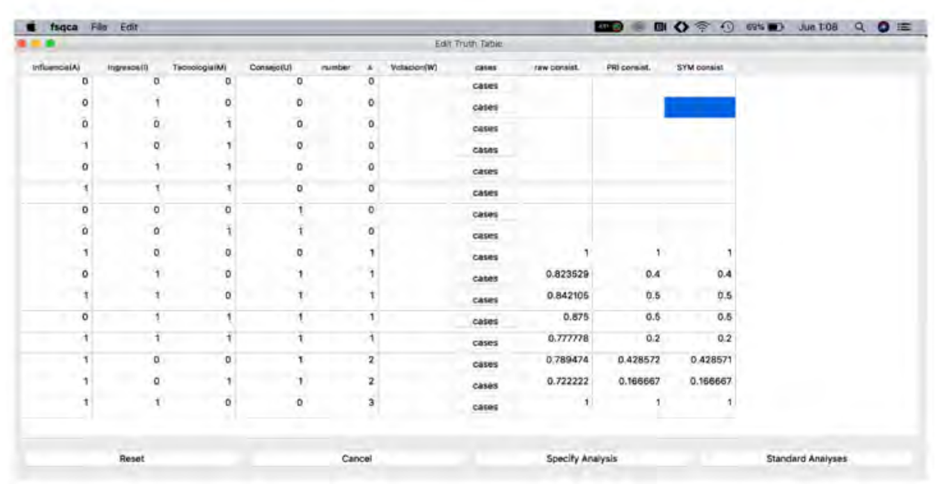
## Ajustando el resultado (*Outcome*). Caso 6

Tenga en cuenta, que la columna etiquetada como el **Outcome** (*Calidad de votación por liderazgo tecnológico*, **W** en este ejemplo) está en blanco. Depende del investigador, determinar el resultado para cada configuración e ingresarlo en la hoja de cálculo utilizando el siguiente procedimiento.

1. El investigador, debe comenzar desarrollando una regla para clasificar algunas combinaciones (*esquinas del espacio vectorial*), como relevantes y otras como *irrelevantes,* basados en el número de casos que se encuentran en cada sector del espacio vectorial, el cual es definido por las *condiciones causales*. Esto se logra, seleccionando un umbral de frecuencia basado en el número de casos más grandes a **0.5** de pertenencia en cada configuración, como se muestra en la columna **number**. Cuando el número total de casos **N**, en un análisis es relativamente pequeño, el *umbral de frecuencia* debe ser **1** o **2**. Cuando **N** es grande, sin embargo, debe usar un umbral más sustancial. Es muy importante examinar la distribución de casos, entre las combinaciones causales, para identificar los sectores más poblados del espacio vectorial. En general, las *configuraciones* seleccionadas deben capturar al menos el **75-80%** de los casos.
2. Los casos se deben ordenar por su frecuencia (*ascendente o descendente*), oprimiendo en el encabezado de la columna **number**.



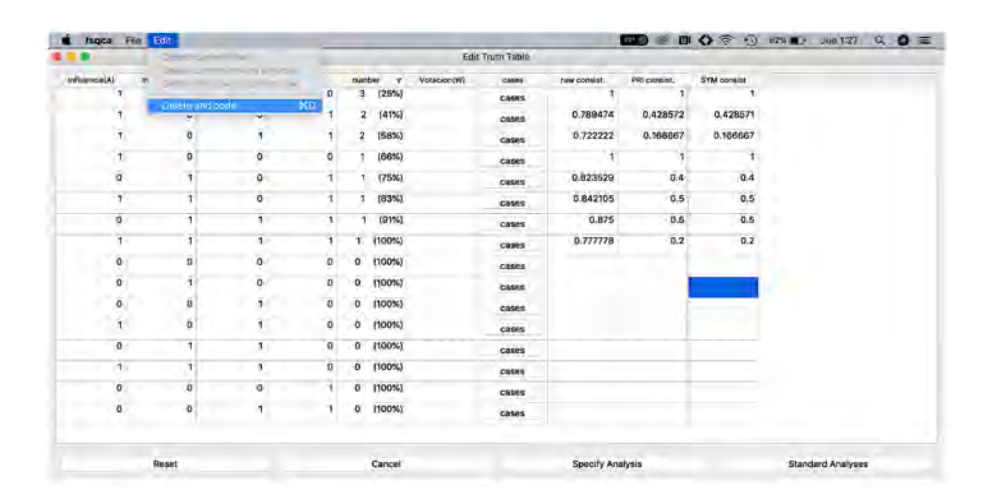

| Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 3 (25%) | | cases | 1 | 1 | 1 |
| 1 | 0 | 0 | 1 | 2 (41%) | | cases | 0.789474 | 0.428572 | 0.428571 |
| 1 | 0 | 1 | 1 | 2 (58%) | | cases | 0.722222 | 0.166667 | 0.166667 |
| 1 | 0 | 0 | 0 | 1 (66%) | | cases | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 (75%) | | cases | 0.823529 | 0.4 | 0.4 |
| 1 | 1 | 0 | 1 | 1 (83%) | | cases | 0.842105 | 0.5 | 0.5 |
| 0 | 1 | 1 | 1 | 1 (91%) | | cases | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 1 | 1 | 1 (100%) | | cases | 0.777778 | 0.2 | 0.2 |
| 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 0 | 0 (100%) | | cases | | | |
| 1 | 0 | 1 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 1 | 0 | 0 (100%) | | cases | | | |
| 1 | 1 | 1 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 0 | 1 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 1 | 0 (100%) | | cases | | | |



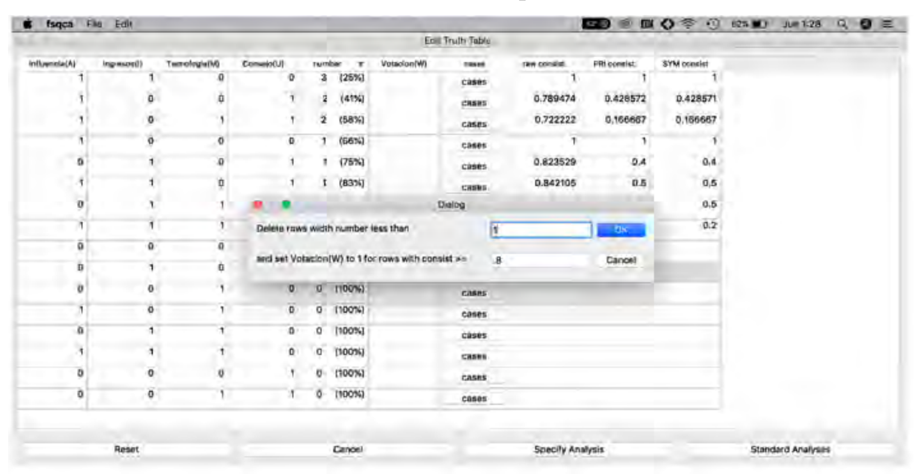

| Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | | cases | | | |
| 0 | 1 | 0 | 0 | 0 | | cases | | | |
| 0 | 0 | 1 | 0 | 0 | | cases | | | |
| 1 | 0 | 1 | 0 | 0 | | cases | | | |
| 0 | 1 | 1 | 0 | 0 | | cases | | | |
| 1 | 1 | 1 | 0 | 0 | | cases | | | |
| 0 | 0 | 0 | 1 | 0 | | cases | | | |
| 0 | 0 | 1 | 1 | 0 | | cases | | | |
| 1 | 0 | 0 | 0 | 1 | | cases | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 | | cases | 0.823529 | 0.4 | 0.4 |
| 1 | 1 | 0 | 1 | 1 | | cases | 0.842105 | 0.5 | 0.5 |
| 0 | 1 | 1 | 1 | 1 | | cases | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 1 | 1 | 1 | | cases | 0.777778 | 0.2 | 0.2 |
| 1 | 0 | 0 | 1 | 2 | | cases | 0.789474 | 0.428572 | 0.428571 |
| 1 | 0 | 1 | 1 | 2 | | cases | 0.722222 | 0.166667 | 0.166667 |
| 1 | 1 | 0 | 0 | 3 | | cases | 1 | 1 | 1 |

3. Después de ordenar y seleccionar un umbral de frecuencia, elimine todas las filas que no cumplan el umbral. Si los casos se han ordenado en orden descendente según la columna **number**, oprima en el primer caso que esté por debajo del umbral y posteriormente elija: **Edit → Delete current row to last row.** Si los casos no se han ordenado, entonces, aquellos casos que no cumplan con el umbral se pueden eliminar individualmente, seleccionando la fila que elija: **Edit → Delete current now.**

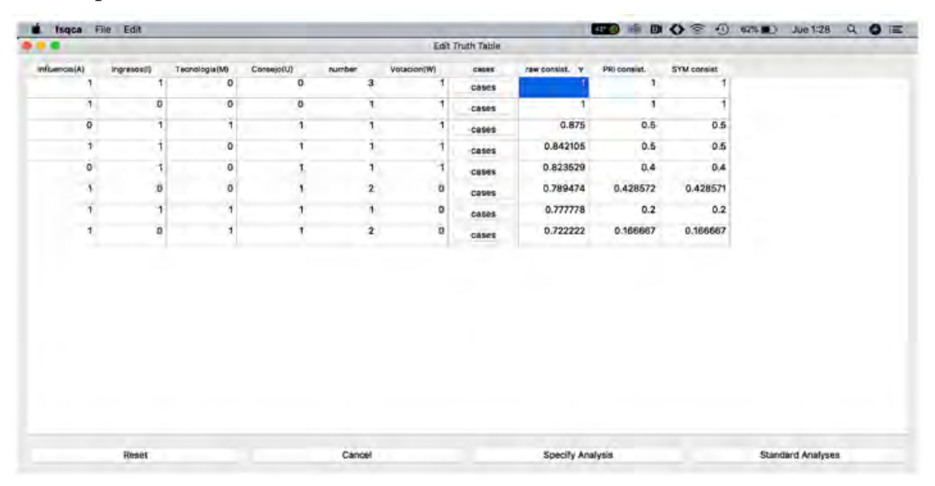
4. El siguiente paso, es distinguir las *configuraciones* que sean subconjuntos consistentes del resultado, de las que no lo son. Esta determinación, se realiza utilizando las medidas de *consistencia* del marco teórico, reportadas en las columnas **raw consist, PRI** y / o **SYM**. Los valores inferiores a **0.80** en la columna **raw consist**, indican una *inconsistencia* sustancial. Es útil ordenar las puntuaciones de *consistencia*, en orden descendente para evaluar su distribución (esto, debe hacerse después de eliminar las filas que no alcanzan el umbral de frecuencia). Ordenar la lista, se logra oprimiendo el encabezado de columna **raw consist**, **PRI** o **SYM** (asegúrese de que la flecha, que aparece al hacer clic en el encabezado de la columna, apunte hacia abajo).
5. Identifique cualquier brecha, en el rango superior de *consistencia*, que pueda ser útil para establecer un umbral de *consistencia*. Tenga en cuenta, que siempre es posible examinar varios umbrales diferentes y evaluar las consecuencias de reducir y/o aumentar, el límite de *consistencia*.
6. Ahora, es necesario indicar qué *configuraciones* exhiben el resultado y cuáles no (ver también, el método alternativo a continuación). Ingrese un (**1**), en la columna de resultados (campo *votación* **W**, en este ejemplo) para cada configuración, cuyo nivel de *consistencia* cumpla o supere el umbral. Ingrese un (**0)**, en la columna de resultados para cada configuración cuyo nivel de *consistencia* no cumpla con el umbral de *consistencia*.
7. Alternativamente, uno puede usar la función **Delete and code** para automatizar este proceso. Seleccione: **Edit → Delete and code**




En el primer campo, se selecciona el umbral de frecuencia. El número predeterminado de casos es **(1)**, pero puede cambiarse escribiendo el umbral de frecuencia seleccionado en el campo. En el segundo campo, se selecciona el umbral de *consistencia* (**raw consist**). La *consistencia* predeterminada es **0.8**, pero esto puede cambiarse escribiendo el umbral de *consistencia* seleccionado en el campo.



Oprima el botón **OK**. El programa eliminará las filas donde no se alcanza el umbral de frecuencia y codificará el resultado como **(0) o (1)** dependiendo del umbral de *consistencia* seleccionado.




| influencia(A) | ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 3 | 1 | cases | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | cases | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | cases | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 0 | 1 | 1 | 1 | cases | 0.842105 | 0.5 | 0.5 |
| 0 | 1 | 0 | 1 | 1 | 1 | cases | 0.823529 | 0.4 | 0.4 |
| 1 | 0 | 0 | 1 | 2 | 0 | cases | 0.789474 | 0.428572 | 0.428571 |
| 1 | 1 | 1 | 1 | 1 | 0 | cases | 0.777778 | 0.2 | 0.2 |
| 1 | 0 | 1 | 1 | 2 | 0 | cases | 0.722222 | 0.166667 | 0.166667 |

**8.** La ventana, muestra la *tabla de verdad* que aparece después de:

**a.** Aplicar un umbral de frecuencia de **(1)**, a los datos y eliminando las *configuraciones* que no tienen observaciones (**8** *configuraciones*).

**b.** Seleccionando un umbral de *consistencia* de **0.8** y colocando un **(1)** en la columna de *votacion* **W,** para *con iguraciones* con *consistencia* de **0.80** o mayor (**5** *configuraciones*) y un **0** para casos con *consistencia* más baja (**3** *configuraciones*).

## Instrucción: Specify Analysis Option.
## Caso 7

A partir de aquí, hay dos posibilidades para el análisis: especificar un análisis único (**Specify Analysis Option**) vs. generar los tres análisis estándar (**Standard Analyses** *complejo, parsimonioso e intermedio*). El procedimiento recomendado a seguir, es oprimir el botón **Standard Analyses** (de la que se recomienda, aplicar la generación de solución intermedia) y explicando a detalle, el funcionamiento de ambas opciones.

Para lograrlo, se debe realizar:

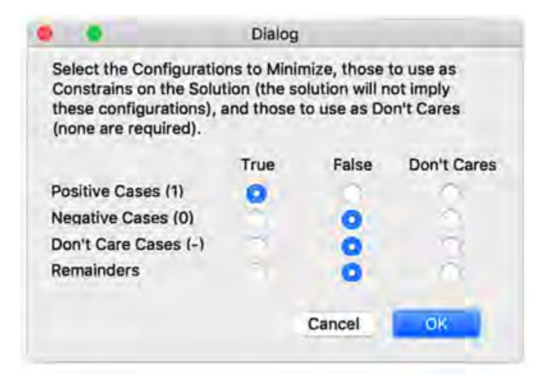
1. Una vez que la *tabla de verdad* ha sido construida, oprima el botón **Specify Analysis** que se ubica en la parte inferior de la ventana:



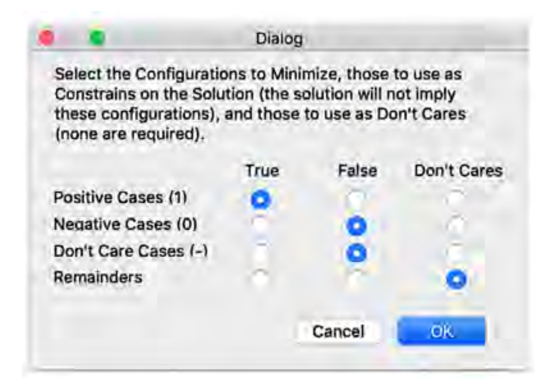

| Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 0 | 1 | 1 | 1 | 0.842105 | 0.5 | 0.5 |
| 0 | 1 | 0 | 1 | 1 | 1 | 0.823529 | 0.4 | 0.4 |
| 1 | 0 | 0 | 1 | 2 | 0 | 0.789474 | 0.428572 | 0.428571 |
| 1 | 1 | 1 | 1 | 1 | 0 | 0.777778 | 0.2 | 0.2 |
| 1 | 0 | 1 | 1 | 2 | 0 | 0.722222 | 0.166667 | 0.166667 |



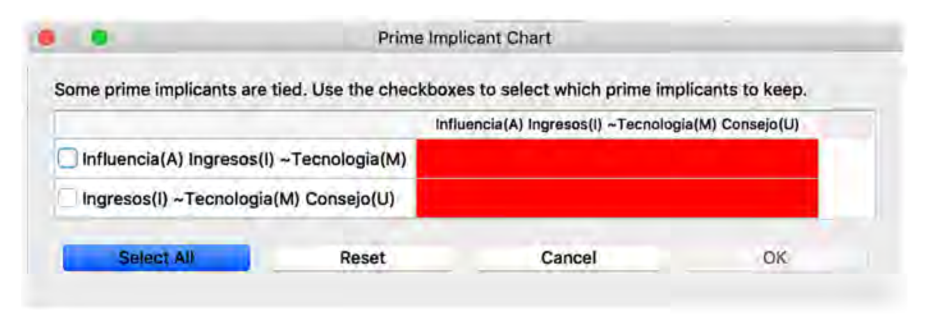
2. En el panel de configuración, marque los casos **Positive Cases** a **True** y todos los demás, en **False**. Esta acción proporcionará la solución más compleja. Esta ventana, aparece como:



3. Para obtener la solución más *parsimoniosa*, establezca **Positive Cases** en **True**, **Negative Cases** a **False** y **Don´t Cares** así como **Remainders** en **Don't Cares:**




4. Tenga en cuenta que, cuando el algoritmo para seleccionar implicaciones principales, no puede reducir completamente la *tabla de verdad*, aparecerá la ventana **Prime Implicant Window** y el usuario debe seleccionar las *implicaciones principales* que se utilizarán, con base al conocimiento del marco teórico de la investigación y el conocimiento empírico sustantivo del investigador. La ventana aparecerá de la siguiente manera:




Las *implicaciones principales*, son términos de producto que se producen utilizando reglas de *minimización* (por ejemplo, las reglas que combinan filas que difieren en una sola causa, si tienen el mismo valor de salida). Por ejemplo: **(ABC)** se combina con **(AbC)** para producir

**(AC). (AC),** por lo tanto, es la implicación principal que cubre las dos expresiones Booleanas primarias **(ABC)** y **(AbC).** En otras palabras, **(ABC)** y **(AbC)** son *subconjuntos* de **(AC),** o **(AC)** implica **(ABC)** y **(AbC).** Sin embargo, a menudo, hay *implicaciones principales* más reducidas que los necesarios, para cubrir todas las expresiones primarias originales y el usuario tiene la opción de elegir entre aquellas que están vinculadas lógicamente, utilizando el gráfico de implicaciones principales **(Ragin, 1987, p. 95)**. Una vez identificadas las combinaciones de condiciones con *el resultado de interés*, **QCA** utiliza la *minimización Booleana* como herramienta para identificar condiciones cuya *presencia o ausenci*a no es relevante para producir un cierto resultado. Según **Marx y Peters (2004, p. 34)**, la minimización sigue la *lógica de un experimento*, de forma que solo se permite que una condición varíe en cada ocasión: si al variar esta condición no observamos un impacto discernible sobre el resultado, podemos decir que era *irrelevante* para su producción. La *minimización Booleana* puede ser entendida como la *reducción de una expresión booleana en una más corta y parsimoniosa* **(Vis, 2010, p. 35)**. Estas combinaciones de condiciones simplificadas se denominan *implicaciones principales*.

**5.** Para elegir las *implicaciones principales*, el programa emplea un algoritmo que intenta reducir la *tabla de verdad,* hasta que no sea posible una mayor simplificación, comenzando con los implicaciones principales esenciales (las que cubren únicamente, filas específicas en la *tabla de verdad*), que deben aparecer en la solución. Si se ejecuta el algoritmo y la tabla no se puede reducir por completo, el usuario puede seleccionar las implicaciones que se utilizarán, en función de su conocimiento teórico y empírica sustantiva**.**

**6.** Aparecerá la ventana **Prime Implicant Chart**, que muestra las posibles implicaciones principales para que el usuario las elija. Cada columna en el gráfico, representa una fila diferente de la *tabla de verdad*, que está cubierta por más de una *implicación principal*.

Prime Implicant Chart

Some prime implicants are tied. Use the checkboxes to select which prime implicants to keep.

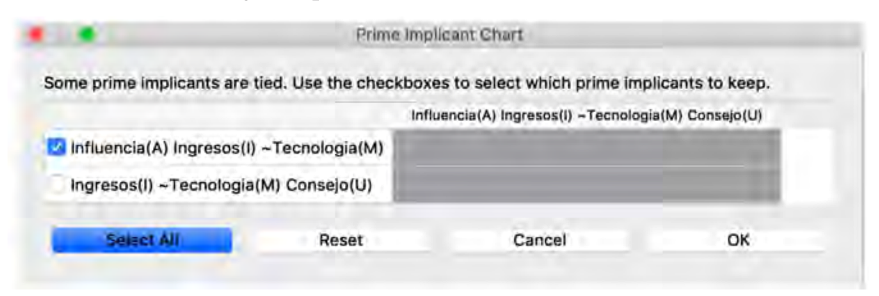
Influencia(A) Ingresos(I) ~Tecnologia(M) Consejo(U)

Influencia(A) Ingresos(I) ~Tecnologia(M)

Ingresos(I) ~Tecnologia(M) Consejo(U)

Select All | Reset | Cancel | OK

La casilla del lado izquierdo describe la *implicación principal* que el usuario puede seleccionar. En nuestro caso, se selecciona la primera fila, como se muestra y se oprime **OK**:




7. Una vez que la *tabla de verdad* se ha *minimizado*, la ventana principal le mostrará el siguiente resultado que representa la solución más compleja. Se debe recordar, que esta solución es la que se obtiene mediante la opción **Specify Analysis**, como se describió anteriormente y bajo las condiciones del **punto 2,** o solución compleja. La primera parte de la salida de resultados, describe los datos del **Truth Table Analysis**. Las dos primeras filas, indican el directorio del archivo y el modelo que especifico. Posteriormente, se muestra qué tipo de algoritmo usó.

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/x/Empresas innovadoras cluster.csv
Model: Votacion(W) = f(Influencia(A), Ingresos(I), Tecnologia(M), Consejo(U))
Algorithm: Quine-McCluskey

--- TRUTH TABLE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.823529
Assumptions:
                                              raw        unique
                                            coverage    coverage
consistency
                                           ----------  ----------
----------
Influencia(A)*~Tecnologia(M)*~Consejo(U)     0.636364    0.030303     1
~Influencia(A)*Ingresos(I)*Consejo(U)        0.515152    0.121212     0.85
Influencia(A)*Ingresos(I)*~Tecnologia(M)     0.636364    0.0303031    0.875
solution coverage: 0.787879
solution consistency: 0.896552


Cases with greater than 0.5 membership in term
Influencia(A)*~Tecnologia(M)*~Consejo(U): D (0.6,0.8),
  E (0.6,0.6), H (0.6,0.8), L (0.6,1)
Cases with greater than 0.5 membership in term
~Influencia(A)*Ingresos(I)*Consejo(U): F (0.6,0.8),
  G (0.6,0.6)
Cases with greater than 0.5 membership in term
Influencia(A)*Ingresos(I)*~Tecnologia(M): A (0.6,0.6),
  D (0.6,0.8), E (0.6,0.6), L (0.6,1)
```

La última parte de la salida de resultados, muestra la **Truth Table Solution,** de su análisis. Primero, se enumeran los límites de frecuencia y *consistencia*. Esto es seguido por la(s) solución(es). Cuando se selecciona la opción **Specify Analysis**, habrá una sección de **Thruth Table Solution**. Cuando se selecciona **Standard Analyses**, se proporcionarán tres secciones de solución (*compleja, parsimoniosa e intermedia*). En este ejemplo, se reportó la solución más compleja, con *calidad de votación* **W** por liderazgo tecnológico de la empresa clúster**,** como un producto de tres *configuraciones*. Las primeras dos líneas, muestran los límites de **frequency** y **consistency cutoff**. Este último, mostrará el valor de *consistencia* más bajo por encima del valor de corte (**consistency cutoff**) especificado por el usuario. Aquí, **0.8** se dio como el límite de *consistencia* (**consistency cutoff**), y el valor real más bajo por encima de **0.8** fue **0.823529.**

**8.** A continuación, se muestra el resultado para la configuración descrita en el **punto 3,** o solución *parsimoniosa*.

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/x/Empresas innovadoras cluster.csv
Model: Votacion(W) = f(Influencia(A), Ingresos(I), Tecnologia(M), Consejo(U))
Algorithm: Quine-McCluskey

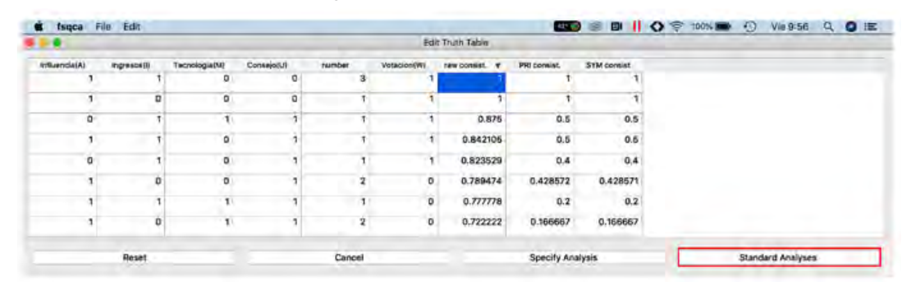
--- TRUTH TABLE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.823529
Assumptions:
                                 raw        unique
                              coverage     coverage    consistency
                             ----------   ----------   ----------
~Consejo(U)                   0.727273     0.0606061    1
~Influencia(A)                0.606061     0.121212     0.869565
Ingresos(I)*~Tecnologia(M)    0.69697      0.030303     0.851852
solution coverage: 0.909091
solution consistency: 0.882353
```

**9.** La solución *parsimoniosa*, proporciona una línea, para cada ruta separada al resultado (en este ejemplo, existen tres rutas). El resultado, también calcula la *consistencia* y la *cobertura* para cada término de solución y la solución en su conjunto. La solución indica, que existen tres trayectorias que logran una baja *calidad de votación* **W**, por liderazgo tecnológico (recuerde que solo **1/12** tiene calidad máxima = **1**), siendo empresas innovadoras que provengan del clúster con pertenencia:

- Medias/altas fortalezas en los resultados del consejo de administración **U**, o
- Baja/nula influencia dentro del clúster **(a)** con altas diferencias de ingresos por innovación **(i)** y que tienen un bajo/nulo grado de tecnología de manufactura **(m).**

## Instrucción: Standard Analyses Option. Caso 8

Una vez que la *tabla de verdad* esté completamente construída (**caso 6**), seleccione **Standard Analyses:**

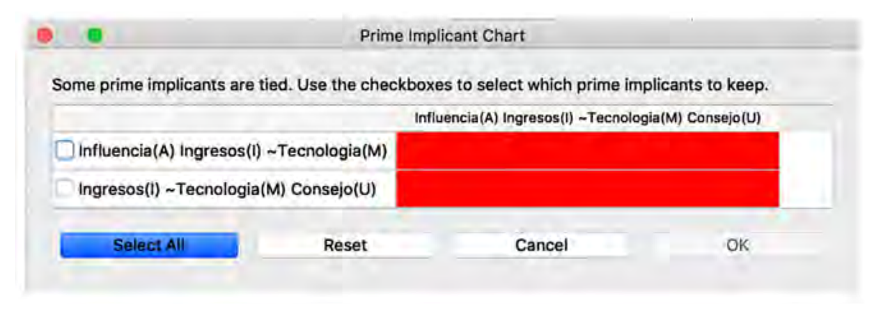



| Influencia(A) | Ingresos(I) | Tecnologia(M) | Consejo(U) | number | Votacion(W) | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 0.875 | 0.5 | 0.5 |
| 1 | 1 | 0 | 1 | 1 | 1 | 0.842105 | 0.5 | 0.6 |
| 0 | 1 | 0 | 1 | 1 | 1 | 0.823529 | 0.4 | 0.4 |
| 1 | 0 | 0 | 1 | 2 | 0 | 0.789474 | 0.428572 | 0.428571 |
| 1 | 1 | 1 | 1 | 1 | 0 | 0.777778 | 0.2 | 0.2 |
| 1 | 0 | 1 | 1 | 2 | 0 | 0.722222 | 0.166667 | 0.166667 |



1. Esta opción, proporciona de forma automática al usuario, soluciones complejas, *parsimoniosas* e intermedias, recordando su uso:
   - **a.** *Solución compleja*: los *residuales* están configurados todos en falso; sin *contrafactuales.*
   - **b.** *Solución parsimoniosa:* es utilizado cualquier *residual*, que ayude a generar la solución más simple, independientemente de si constituye un *contrafactual* fácil o difícil.
   - **c.** *Solución intermedia:* Este es el procedimiento recomendado a utilizar, ya que es la única forma, por cierto, de obtener la solución intermedia. Para generarla, el software realiza análisis *contrafactuales*, basados en información sobre las *condiciones causales* que proporciona directamente el investigador. Solo se permiten los *residuales* que son casos *contrafactuales fáciles* para ser incorporados a la solución. La designación de fácil vs. difícil se basa en información proporcionada por el investigador, con respecto a la conexión entre cada *condición causal* y el resultado.
2. Tenga en cuenta que, cuando el algoritmo para seleccionar implicaciones principales, no puede reducir completamente la *tabla de verdad*, aparecerá la ventana **Prime Implicant Window** y el usuario debe seleccionar los implicaciones principales que se utilizarán, con base al conocimiento

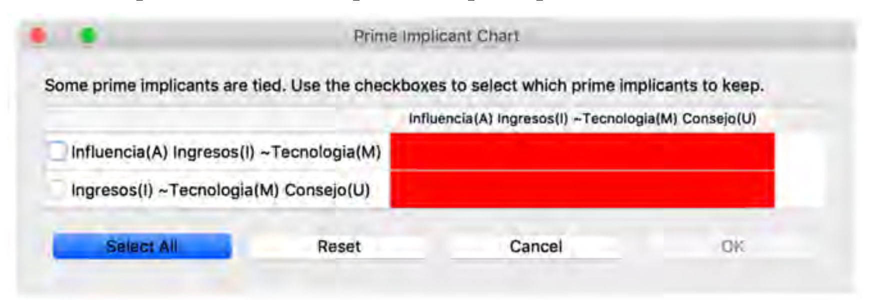
del marco teórico de la investigación y el conocimiento empírico sustantivo del investigador. La ventana aparecerá de la siguiente manera:



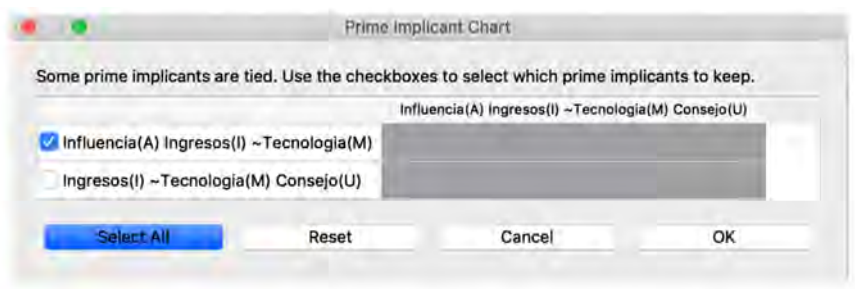

Las implicaciones principales, son términos de producto que se producen utilizando reglas de minimización (por ejemplo, las reglas que combinan filas que difieren en una sola causa, si tienen el mismo valor de salida). Por ejemplo: **(ABC)** se combina con **(AbC)** para producir **(AC). (AC),** por lo tanto, es la implicación principal que cubre las dos expresiones Booleanas primarias **(ABC)** y **(AbC).** En otras palabras, **(ABC)** y (**AbC**) son subconjuntos de **(AC),** o **(AC)** implica **(ABC)** y **(AbC).** Sin embargo, a menudo, hay implicaciones principales más reducidos que los necesarios, para cubrir todas las expresiones primarias originales y el usuario tiene la opción de elegir entre aquellas que están vinculadas lógicamente, utilizando el gráfico de implicaciones principales **(Ragin, 1987)**.

3. Para elegir las implicaciones principales, el programa emplea un algoritmo que intenta reducir la *tabla de verdad*, hasta que no sea posible una mayor simplificación, comenzando con los implicaciones principales esenciales (las que cubren únicamente, filas específicas en la *tabla de verdad*), que deben aparecer en la solución. Si se ejecuta el algoritmo y la tabla no se puede reducir por completo, el usuario puede seleccionar las implicaciones que se utilizarán, en función de su conocimiento teórico y empírica sustantiva**.**
4. Aparecerá la ventana **Prime Implicant Chart**, que muestra las posibles implicaciones principales para que el usuario las elija. Cada columna en

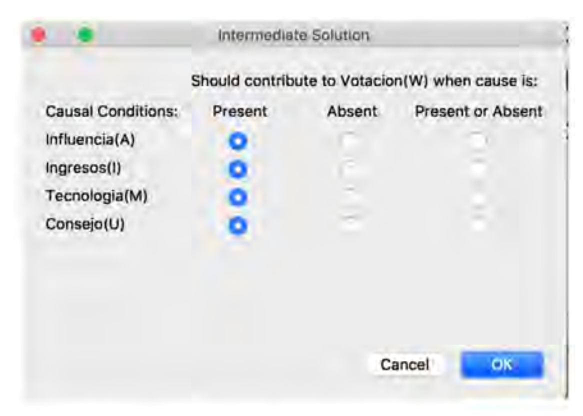
el gráfico, representa una fila diferente de la *tabla de verdad*, que está cubierta por más de una implicación principal.



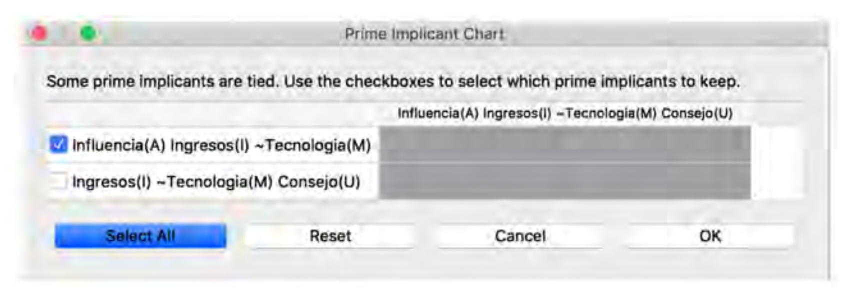

La casilla del lado izquierdo describe la implicación principal que el usuario puede seleccionar. En nuestro caso, se selecciona la primera fila, como se muestra y se oprime **OK**:




5. Para generar la solución intermedia, aparecerá la ventana **Intermediate Solution** de la que se sugiere marcar todas las variables o *condiciones causales* en **Present:**



6. Nuevamente, puede aparecer la pantalla **Prime Implicant Chart,** que se sugiere marcar como se había hecho en el punto 4 arriba mencionado:




7. De esta manera, se obtienen los listados de la solución compleja y *parsimoniosa*, mencionados en los **puntos 7 y 8** de **Specify Analysis,** obteniendo:

   **a.** La solución **compleja**:

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/x/Empresas innovadoras cluster.csv
Model: Votacion(W) = f(Influencia(A), Ingresos(I), Tecnologia(M), Consejo(U))
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.823529
                                                  raw       unique
                                               coverage    coverage
consistency
                                              ----------  ----------
----------
Influencia(A)*~Tecnologia(M)*~Consejo(U)      0.636364    0.030303    1
~Influencia(A)*Ingresos(I)*Consejo(U)         0.515152    0.121212    0.85
Influencia(A)*Ingresos(I)*~Tecnologia(M)      0.636364    0.0303031   0.875
solution coverage: 0.787879
solution consistency: 0.896552


Cases with greater than 0.5 membership in term
Influencia(A)*~Tecnologia(M)*~Consejo(U): D (0.6,0.8),
  E (0.6,0.6), H (0.6,0.8), L (0.6,1)
Cases with greater than 0.5 membership in term
~Influencia(A)*Ingresos(I)*Consejo(U): F (0.6,0.8),
  G (0.6,0.6)
Cases with greater than 0.5 membership in term
Influencia(A)*Ingresos(I)*~Tecnologia(M): A (0.6,0.6),
  D (0.6,0.8), E (0.6,0.6), L (0.6,1)
```

**b.** La solución **parsimoniosa:**

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/x/Empresas innovadoras cluster.csv
Model: Votacion(W) = f(Influencia(A), Ingresos(I), Tecnologia(M), Consejo(U))
Algorithm: Quine-McCluskey

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.823529
                                  raw       unique
                                coverage   coverage   consistency
                               ---------- ---------- ----------
~Consejo(U)                     0.727273   0.0606061  1
~Influencia(A)                  0.606061   0.121212   0.869565
Ingresos(I)*~Tecnologia(M)      0.69697    0.030303   0.851852
solution coverage: 0.909091
solution consistency: 0.882353


Cases with greater than 0.5 membership in term ~Consejo(U): D (0.8,0.8),
  L (0.8,1), E (0.6,0.6), H (0.6,0.8)
Cases with greater than 0.5 membership in term ~Influencia(A): F (0.8,0.8),
  G (0.6,0.6)
Cases with greater than 0.5 membership in term Ingresos(I)*~Tecnologia(M): D
(0.8,0.8),
  G (0.8,0.6), A (0.6,0.6), E (0.6,0.6),
  L (0.6,1)
```

**c.** Con el adicional de la solución **intermedia**, como se muestra:

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/x/Empresas innovadoras cluster.csv
Model: Votacion(W) = f(Influencia(A), Ingresos(I), Tecnologia(M), Consejo(U))
Algorithm: Quine-McCluskey

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.823529
Assumptions:
Influencia(A) (present)
Ingresos(I) (present)
Tecnologia(M) (present)
Consejo(U) (present)
                                               raw        unique
                                             coverage    coverage
consistency
                                            ----------  ----------
----------
Influencia(A)*~Consejo(U)                    0.69697     0.0606061   1
~Influencia(A)*Ingresos(I)*Consejo(U)        0.515152    0.0909091   0.85
Influencia(A)*Ingresos(I)*~Tecnologia(M)     0.636364    0.0303031   0.875
solution coverage: 0.818182
solution consistency: 0.9


Cases with greater than 0.5 membership in term Influencia(A)*~Consejo(U): L
(0.8,1),
  D (0.6,0.8), E (0.6,0.6), H (0.6,0.8)
Cases with greater than 0.5 membership in term
~Influencia(A)*Ingresos(I)*Consejo(U): F (0.6,0.8),
  G (0.6,0.6)
Cases with greater than 0.5 membership in term
Influencia(A)*Ingresos(I)*~Tecnologia(M): A (0.6,0.6),
  D (0.6,0.8), E (0.6,0.6), L (0.6,1)
```

8. La parte de **Assumptions** (supuestos) de los resultados de salida, muestra las opciones previamente seleccionadas en la ventana **Intermediate Solution**; aquí, cada uno fue seleccionado de tal manera que, cuando está presente, la condición debería contribuir al resultado.
9. La solución indica, que existen tres trayectorias que logran una baja *calidad de votación por liderazgo tecnológico* **W**, (recuerde que solo **1/12** tiene calidad máxima = **1**), siendo empresas innovadoras que provengan del clúster con pertenencia:
   a. Media/alta i*nfluencia dentro del clúster* **A** pero con nula/baja fortaleza en los *resultados del consejo de administración* **(u).**
   b. Baja/nula *influencia dentro del clúster* **(a),** con medias/altas *diferencias por ingresos de innovación* **I** y con medias/altas *fortalezas en los resultados del consejo de administració*n **U.**
   c. Medias/altas *in luencias dentro del clúster* **A**, con medias/altas *diferencias por ingresos de innovación* **I**, pero con bajo/nulo *grado de tecnología en manufactura* (**m**).

## Preparando datos difusos para calcular la consistencia y cobertura. Caso 9

El resultado, incluye medidas de *cobertura* y *consistencia* para cada término de solución y para la solución como un todo. La *consistencia* (con la suficiencia), miden el grado en que los términos de la solución y la solución en su conjunto, son subconjuntos del resultado. La *cobertura*, mide cuánto del resultado está cubierto (o explicado), por cada término de solución y por la solución como un todo. Estas medidas, se calculan examinando los conjuntos de *datos difusos* originales, a la luz de la solución (compuesto por uno o más términos de solución). El grado, en que los casos en el conjunto de datos original, tienen pertenencia en cada término de la solución y en el resultado, forma la base de las medidas de *consistencia* y *cobertura*.

Más específicamente, considere la siguiente **Tabla 4.19** de datos hipotéticos, con tres *condiciones causales* **(A, B y C)** y un resultado **Y,** todos medidos como conjuntos de *datos difusos*.

**Tabla 4.19. Tabla de datos hipotéticos para prueba de *consistencia* y *cobertura***

| Item | Columnas | | | |
|---|---|---|---|---|
| | Pertenencia de condiciones causales | | | Pertenencia de resultado |
| | A | B | C | Y |
| 1 | 0.8 | 0.9 | 0.8 | 0.9 |
| 2 | 0.6 | 0.7 | 0.4 | 0.8 |
| 3 | 0.6 | 0.7 | 0.2 | 0.7 |
| 4 | 0.6 | 0.6 | 0.3 | 0.7 |
| 5 | 0.8 | 0.3 | 0.7 | 0.8 |
| 6 | 0.6 | 0.1 | 0.7 | 0.9 |
| 7 | 0.7 | 0.4 | 0.2 | 0.3 |
| 8 | 0.2 | 0.9 | 0.9 | 0.1 |
| 9 | 0.1 | 0.6 | 0.2 | 0.2 |
| 10 | 0.2 | 0.1 | 0.7 | 0.3 |
| 11 | 0.3 | 0.1 | 0.3 | 0.3 |
| 12 | 0.1 | 0.2 | 0.3 | 0.2 |
| **Total** | | | | **6.2** |

Fuente: Ragin (2008) con adaptación propia.

Realice las diferentes combinaciones causales, así como los cálculos para comprobar el **raw coverage, unique coverage y consistency**, que reporta el resultado final, recordando que el **AND (*)** lógico implica la **función mín** y el **OR (+)** lógico la **función máx** del valor de los grupos en comparación.

La relevancia de este análisis, se muestra a continuación:

1. La solución se compone de dos términos: (**A * B**) + (**A * C**). Para calcular la *consistencia* y la *cobertura*, primero se deben calcular varios valores intermedios.
2. La pertenencia en el resultado, $\Sigma_y$, es la suma de las puntuaciones de pertenencia a través de todos datos de los casos.

**3.** La pertenencia de un caso en cada término de la solución, se calcula como el **AND** lógico que implica la función **mín.** del valor de la pertenencia de los casos, en cada *condición causal* del término.

**4.** La pertenencia, en el primer término de solución $\Sigma_{A*B}$, es la suma de la pertenencia en ese término, de la solución en todos los casos.

**5.** De manera similar, la membresía en el segundo término de solución ($\Sigma_{A*C}$), es la suma de la pertenencia en ese término de solución en todos los casos.

**6.** La pertenencia en la solución ($\Sigma_{(A*B+A*C)}$), es definida con el **OR** lógico que implica la función **máx.** del valor de la pertenencia de los casos, en cada *condición causal* del término. Ver **Tabla 4.20.**

**Tabla 4.20. Tabla de datos hipotéticos para prueba de consistencia y *cobertura* con datos difusos intermedios de pertenencia resultado vs. solución**

| Item | Columnas | | | | | | |
|---|---|---|---|---|---|---|---|
| | Pertenencia de condiciones causales | | | Pertenencia de resultado | Pertenencia de la solución | | |
| | A | B | C | Y | (A * B) | (A * C) | (A * B) + (A * C) |
| 1 | 0.8 | 0.9 | 0.8 | 0.9 | 0.8 | 0.8 | 0.8 |
| 2 | 0.6 | 0.7 | 0.4 | 0.8 | 0.6 | 0.4 | 0.6 |
| 3 | 0.6 | 0.7 | 0.2 | 0.7 | 0.6 | 0.2 | 0.6 |
| 4 | 0.6 | 0.6 | 0.3 | 0.7 | 0.6 | 0.3 | 0.6 |
| 5 | 0.8 | 0.3 | 0.7 | 0.8 | 0.3 | 0.7 | 0.7 |
| 6 | 0.6 | 0.1 | 0.7 | 0.9 | 0.1 | 0.6 | 0.6 |
| 7 | 0.7 | 0.4 | 0.2 | 0.3 | 0.4 | 0.2 | 0.4 |
| 8 | 0.2 | 0.9 | 0.9 | 0.1 | 0.2 | 0.2 | 0.2 |
| 9 | 0.1 | 0.6 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 |
| 10 | 0.2 | 0.1 | 0.7 | 0.3 | 0.1 | 0.2 | 0.2 |
| 11 | 0.3 | 0.1 | 0.3 | 0.3 | 0.1 | 0.3 | 0.3 |
| 12 | 0.1 | 0.2 | 0.3 | 0.2 | 0.1 | 0.1 | 0.1 |
| **Total** | | | | **6.2** | **4.0** | **4.1** | **5.2** |

Fuente: Ragin (2008) con adaptación propia.

## Cálculo de la consistencia (*consistency*). Caso 10

Con los preparativos anteriores, y partiendo de la **Tabla 4.20** es factible, calcular tanto los factores de *consistencia* (*consistency*), como de solución por *consistencia* (*solution consistency*). Las medidas de *consistencia y cobertura* son dos conceptos esenciales para entender cualquier solución **QCA** tras la *minimización Booleana*. Ver **Tabla 4.21.**

**Tabla 4.21. Tabla de datos hipotéticos para cálculo de *consistencia* y *cobertura***

| Item | Columna | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Pertenencia de condiciones causales | | | Pertenencia de resultado | Pertenencia de la solución | | | Cálculos de consistencia | | |
| | A | B | C | Y | (A * B) | (A * C) | (A * B) + (A * C) | 1 | 2 | 3 |
| 1 | 0.8 | 0.9 | 0.8 | 0.9 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 |
| 2 | 0.6 | 0.7 | 0.4 | 0.8 | 0.6 | 0.4 | 0.6 | 0.6 | 0.4 | 0.6 |
| 3 | 0.6 | 0.7 | 0.2 | 0.7 | 0.6 | 0.2 | 0.6 | 0.6 | 0.2 | 0.6 |
| 4 | 0.6 | 0.6 | 0.3 | 0.7 | 0.6 | 0.3 | 0.6 | 0.6 | 0.3 | 0.6 |
| 5 | 0.8 | 0.3 | 0.7 | 0.8 | 0.3 | 0.7 | 0.7 | 0.3 | 0.7 | 0.7 |
| 6 | 0.6 | 0.1 | 0.7 | 0.9 | 0.1 | 0.6 | 0.6 | 0.1 | 0.6 | 0.6 |
| 7 | 0.7 | 0.4 | 0.2 | 0.3 | 0.4 | 0.2 | 0.4 | 0.3 | 0.2 | 0.3 |
| 8 | 0.2 | 0.9 | 0.9 | 0.1 | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 |
| 9 | 0.1 | 0.6 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| 10 | 0.2 | 0.1 | 0.7 | 0.3 | 0.1 | 0.2 | 0.2 | 0.1 | 0.2 | 0.2 |
| 11 | 0.3 | 0.1 | 0.3 | 0.3 | 0.1 | 0.3 | 0.3 | 0.1 | 0.3 | 0.3 |
| 12 | 0.1 | 0.2 | 0.3 | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| Σ | | | | 6.2 | 4.0 | 4.1 | 5.2 | 3.8 | 4.0 | 5.0 |

1. A*B ; 2. A*C; 3. A*B+A*C

Fuente: Ragin (2008) con adaptación propia.

La *consistencia* especifica cuántos de los casos explicados por la solución son en realidad casos donde el resultado de interés puede ser observado, mientras que la *cobertura* explica cuántos de los casos con el resultado de interés están cubiertos por la solución **(Rihoux y De Meur, 2009,**

**pp. 47 y 64)**. Por utilizar una expresión coloquial, la *cobertura explica si están todos los que son*, mientras la *consistencia explica si son todos los que están*.

La *consistencia* (*consistency*), mide el grado en el cual, la pertenencia por cada término de solución, es un subconjunto del resultado. Por lo que:

1. La *consistencia*, se calcula primero por cada caso componente. Para cualquier término de la solución, un caso es consistente, si la pertenencia en el término de la solución, es menor o igual que la pertenencia en el resultado.(Comparar columna **Y** vs. **A*B** que el menor, producirá la columna $\boldsymbol{C_{A*B}}$).
2. Si la pertenencia de un caso, en el término de la solución, es mayor que su pertenencia en el resultado (es decir, es inconsistente), entonces el caso recibe una puntuación que equivale a su pertenencia en el resultado.
3. Estos puntajes se suman (generando $\boldsymbol{\Sigma C_{A*B}}$), y se dividen, por la suma de las pertenencias en el término de la solución $\boldsymbol{\Sigma_{A*B}}$).
4. Por lo tanto, la *consistencia* (*consistency*), para el primer término de solución es:

$$(\Sigma C_{A*B} \;/\; \Sigma_{A*B} = 3.8 \;/\; 4.0 = 0.950000$$

$$(\Sigma C_{A*C} \;/\; \Sigma_{A*C} = 4.0 \;/\; 4.1 = 0.975610$$

Así, la *consistencia* (*consistency*) en este ejemplo de (**A * B**), es **0.950000** y de (**A * C**), es **0.975610**. Ver **Tabla 4.22**

**Tabla 4.22. Resultados directos de fsQCA**

| | raw coverage | unique coverage | consistency |
|---|---|---|---|
| A*B+ | 0.612903 | 0.161290 | 0.950000 |
| A*C | 0.645161 | 0.193548 | 0.975610 |
| solution coverage: 0.806452 | | | |
| solution consistency: 0.961538 | | | |

Fuente: Elaboración propia.

## Cálculo de la solución por consistencia (*solution consistency*). Caso 11

El cálculo de la solución por *consistencia* (*solution consistency*) mide el grado, en que la pertenencia en la solución (el conjunto de términos de la solución), es un subconjunto de pertenencia, en el resultado. Por lo que:

1. El máximo de la pertenencia de cada caso, a través de los términos de la solución **max** (**A * B + A * C**), se compara con la pertenencia en el resultado.
2. Si la pertenencia en la solución es, menor o igual que la pertenencia en el resultado, entonces el caso recibe una puntuación que es igual a su pertenencia en el término de la solución.
3. Si la pertenecia en el término de la solución, es mayor que la pertenecia en el resultado (es decir, si es inconsistente) entonces, el caso recibe las puntuaciones del resultado (la más baja de las dos puntuaciones).
4. Estas puntuaciones se suman y luego se dividen por la suma de las pertenencias en el término de la solución:

$$(\Sigma C_{(A*B+A*C)} / \Sigma_{(A*B+A*C)} = 5.0 / 5.2 = 0.961538$$

Así, la solución por *consistencia* (*solution consistency*) en este ejemplo es **0.961538.** Ver **Tabla 4.22.**

## Cálculo de solución por cobertura (*solution coverage*). Caso 12

Con los preparativos anteriores, y partiendo de la **Tabla 4.20,** es factible calcular tanto los factores de *solución por cobertura* (*solution coverage*), *cobertura bruta* (*raw coverage*) y *cobertura única* (*unique coverage*).

El cálculo de *solución por cobertura* (*solution coverage*), mide la proporción de pertenencias en el resultado, que se explica por la solución completa. Las puntuaciones de pertenencia consistentes, se suman entre los casos y posteriormente, se dividen por la suma de las pertenencias en el resultado:

$$(\Sigma C_{(A*B+A*C)} / \Sigma Y = 5.0 / 6.2 = 0.806452$$

Así, la *solución por cobertura* (*solution coverage*) en este ejemplo es **0.806452.** Ver **Tabla 4.22.**

## Cálculo de cobertura bruta (*raw coverage*). Caso 13

El cálculo de *cobertura bruta* (*raw coverage*), mide la proporción de pertenencias en el resultado, que se explica por cada término de la solución. El cálculo de *cobertura bruta* (*raw coverage*), se calcula por cada término de la solución, de los datos originales, al dividir la suma de las pertenencias consistentes en el término de la solución, por la suma de las pertenencias en el resultado:

$$(\Sigma C_{(A*B)} / \Sigma Y = 3.8 / 6.2 = 0.612903$$

$$(\Sigma C_{(A*C)} / \Sigma Y = 4.0 / 6.2 = 0.645161$$

Así, la *solución por cobertura bruta* (*raw coverage*) en este ejemplo, para (**A*B**), es **0.612903** y de (**A*C**), es **0.645161.** Ver **Tabla 4.22.**

## Cálculo de cobertura única (*unique coverage*). Caso 14

El cálculo de *cobertura única* (*unique coverage*), mide la proporción de pertenencias, en el resultado explicado únicamente por cada término de solución individual (pertenencias que no están cubiertas, por otros términos de solución). Por lo que:

**1.** Esto se realiza, eliminando primero el término de la solución y calculando la *solución por cobertura* (*solution coverage*).

**2.** En este ejemplo, la *solución por cobertura* (*solution coverage*), después de eliminar el primer término de solución $\Sigma C_{(A*B)}$ es simplemente

$\Sigma C_{(A*C)}$ (con **n** términos de solución, la solución reducida contendrá los términos de solución **n-1**).

3. El término de *cobertura* reducida (*reduced coverage term*), se divide por la *solución por cobertura* (*solution coverage*) completa y se resta de la *cobertura bruta* (*raw coverage*) para proporcionar la *cobertura única* (*unique coverage*), para el término de solución omitido (*omitted solution term*).
4. Para el primer término de solución $\Sigma C_{(A*B)}$, la *cobertura única* (*unique coverage*), es igual a:

$$(\Sigma C_{(A*B+A*C)} / \Sigma Y) - (\Sigma C_{(A*C)} / \Sigma Y) = (5.0 / 6.2) - (4.0 / 6.2) = 0.161290$$

5. Para el segundo término de solución $\Sigma C_{(A*B)}$, la *cobertura única* (*unique coverage*), es igual a:

$$(\Sigma C_{(A*B+A*C)} / \Sigma Y) - (\Sigma C_{(A*B)} / \Sigma Y) = (5.0 / 6.2) - (3.8 / 6.2) = 0.193548$$

Ver **Tabla 4.22.**

## Resumen fsQCA para la construcción de conjuntos de datos difusos (*fuzzy-sets*)

Más adelante, se presentarán dos medidas descriptivas simples para evaluar las conexiones basadas en el marco teórico de la investigación, como la que se muestra en la **Gráfica 4.1**. Específicamente, se demostrará que es posible evaluar:

1. El grado en que la evidencia empírica es consistente con la afirmación de que un existe una conexión con la propuesta del marco teórico de la investigación, y
2. La importancia o relevancia empírica de tal conexión.

Los conjuntos de *datos difusos*, son fáciles de detectar utilizando gráficas simples **XY** de difusos de puntajes de pertenencia. En general, diagramas triangulares, con puntos consistentemente arriba o consistentemente debajo

de cualquier diagonal de la gráfica, que señala un subconjunto de *datos difusos*, de algún tipo. El análisis **fsQCA**, incluye facilidades para graficar los conjuntos de *datos difusos*. Ver **Tabla 4.23.**

**Tabla 4.23. Resumen fsQCA para la construcción del conjunto de datos difusos (*fuzzy-sets*)**

| Item | Etapa | Descripción |
|---|---|---|
| **1** | **Creando el conjunto de datos** | Esta etapa se puede lograr:<br>**a.** Usando **fsQCA**, que incluye procedimientos para ingresar datos directamente y para importar conjuntos de datos de otros programas (por ejemplo, archivos delimitados por comas desde Excel o archivos delimitados por tabulaciones desde **SPSS**).<br>**b.** Los archivos importados deben tener nombres de variables simples (sin espacios incrustados o puntuación) como la primera línea del conjunto de datos.<br>**c.** Los datos faltantes deben ingresarse como espacios en blanco. El puntaje de pertenencia difusa puede ser asignado o calculado utilizando los procedimientos detallados en los capítulos por venir . |
| **2** | **Ajustes** | Una vez que el conjunto de datos es visible en la ventana de la hoja de cálculo de datos de **fsQCA:**<br>**a.** Haga clic en Gráficos (*Graphs***)**, posteriormente Difuso (*Fuzzy*), posteriormente Graficar en X (*XY Plot*).<br>**b.** Especifique los conjuntos de *datos difusos* (*fuzzy-sets*) para ser graficado en los ejes X e Y haciendo clic en las flechas hacia abajo adyacentes y posteriormente haciendo clic en los nombres de las variables relevantes.<br>**c.** También es posible negar los conjuntos de *datos difusos* (*fuzzy-sets*) antes de graficarlos; haga clic en el cuadro Negar (*Negate*) siguiente al nombre de la variable.<br>**d.** Especifique una variable opcional de caso Id. para que el caso o casos que residen en puntos específicos en la gráfica pueden ser fácilmente identificado. |
| **3** | **Generación de la gráfica** | Examine el patrón y haga clic en cualquier punto de la gráfica y su información aparecerá en la parte inferior de la misma. |
| **4** | **Mostrando valores** | Los números que se muestran en los cuadros sobre la esquina superior izquierda y debajo de la esquina inferior derecha de la gráfica son consistentes a la *cobertura* de los puntajes, que se explicarán más adelante. |

Fuente: Ragin (2008) con adaptación propia.

CAPÍTULO 5.

# Evaluación del conjunto de datos fsQCA

Trataremos en este capítulo, las condiciones **INUS** y **SUIN,** así como las mediciones descriptivas simples para evaluar la fuerza del soporte empírico de los argumentos que sostienen al marco teórico de la investigación de una investigación y sus relaciones.

## Condiciones INUS y SUIN

Un par de conceptos referentes a la parte sustantiva de los análisis **QCA,** que se requiere abordar, son los de condiciones **INUS** y **SUIN**. De hecho, para entenderlos exige un conocimiento previo de los conceptos de *necesidad, suficiencia, condición y configuración*.

Las siglas **INUS** (*lnsufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result*) están relacionadas con el análisis de *suficiencia*, y hacen referencia a condiciones que son partes *insuficientes* pero necesarias de una configuración que es *innecesaria*, pero *suficiente* para producir el *resultado* **(Schneider y Wagemann, 2012, p. 79)**. Una condición **INUS** se caracteriza, pues, por cuatro aspectos:

**a.** Forma parte de una configuración de condiciones que es suficiente para producir el resultado.
**b.** Esa configuración, sin embargo, no es necesaria para producir el resultado, puesto que puede haber otras configuraciones que también conduzcan a él.
**c.** La condición **INUS** no es *suficiente* por sí misma para producir dicha configuración suficiente pero innecesaria, puesto que por definición deberá estar combinada con otras condiciones.

**d.** Sin embargo, al mismo tiempo, la condición **INUS** es necesaria para producir la configuración suficiente pero innecesaria, puesto que, conceptualmente, sin ella la configuración no podría existir.

Por ejemplo, en la expresión **(A*B+C → Y)** la configuración **(A*B)** no es necesaria para producir **Y** puesto que **C** también puede producir el mismo resultado. Sin embargo, esta configuración **(A*B)** sí es *suficiente* para la producción de **Y**. En este ejemplo, tanto **A** como **B** son individualmente necesarios para crear la *configuración suficiente* **(A*B)** aunque individualmente no sean *suficientes* (porque se necesitan la una a la otra) para producir dicha *configuración*. Por tanto, tanto **A** como **B** son condiciones **INUS**.

Por el contrario, las siglas inglesas **SUIN** (*Sufficient, but Unnecessary part of a factor that is Insufficient, but Necessary for the result*) hacen referencia a:

1. Partes *suficientes* pero *innecesarias* de un factor, que es *insuficiente* pero *necesario* para producir el *resultado* **(Schneider y Wagemann, 2012, p. 79)**.
2. Podemos simplificar este concepto como hicimos con el concepto **INUS** anterior, *pero antes debemos tener en cuenta la diferencia entre unión y configuración* **(Schneider y Wagemann, 2012, p. 80)**. Mientras en una *configuración* dos o más condiciones deben producirse simultáneamente **(A*B),** una *unión* hace referencia a dos condiciones que son alternativas entre sí **(A+B).** Reteniendo esto en mente, podemos decir que:
   - **I.** La condición **SUIN** es parte de una unión de condiciones que es necesaria para producir el *resultado*, puesto que el *resultado* nunca se produce sin esa *unión*.
   - **II.** Sin embargo, dicha *unión no es suficiente para producir el resultado*, puesto que debe producirse simultáneamente junto a otra condición o condiciones.
   - **III.** La condición **SUIN** forma parte de una *unión*, por lo que se presenta como alternativa a otra condición dentro de la misma unión. Dado

que ambas condiciones en la unión se presentan como *alternativas con iguales efectos causales, ninguna de ellas es en sí necesaria.*

**IV.** Sin embargo, por el mismo motivo (se presentan como alternativas con iguales efectos) ambas son en sí mismas *suficientes* con independencia de la otra para, en su conjunción con otra condición o condiciones, producir el resultado.

Por ejemplo, en la expresión **(A+B)*C = Y**, vemos que para producir **Y** es necesario que bien **A** o bien **B** se combinen con **C**. En dicha expresión, **A** y **B** son alternativas entre sí, por lo que *ninguna es necesaria*, aunque ambas *son suficientes*, para combinarse con **C** y producir una *configuración necesaria*. Al mismo tiempo, y por lo mismo, la unión **(A+B)** *no es suficiente* para producir **Y**, ya que debe darse simultáneamente junto a **C**. Sin embargo, dicha unión **(A+B)** es *necesaria* para producir el *resultado*, puesto que si ni **A** o ni **B** están presentes el resultado **Y** no podrá producirse, por más que **C** esté presente.

## Evaluando la fuerza del soporte empírico

Basados en **Ragin (2008)**, se tiene:

1. Primero se tratarán los argumentos que afirman que una causa específica o combinación de las *condiciones causales* constituyen uno de los varios caminos posibles, que conducen a un resultado,
2. Cuando lo anterior es cierto, los casos que muestran una combinación causal específica, deben constituir un subconjunto de los casos que muestran el resultado.
3. Así, se presentan medidas para evaluar dos aspectos distintos del marco teórico de la investigación y sus relaciones.

Con lo explicado anteriormente, tenemos que la cobertura del marco teórica de una investigación, por contraste, evalúa el grado en que una causa o combinación causal justifica las instancias que producen un resul-

tado. Cuando hay varios caminos hacia el mismo resultado, la cobertura de una combinación causal dada puede ser pequeña .Por lo tanto, la cobertura mide la relevancia o importancia empírica. Estas mismas medidas, pueden usarse para evaluar situaciones en las que el investigador sospecha que una *condición causal es necesaria* (*pero no suficiente*) para producir un resultado, es decir, donde las instancias de un resultado constituyan un subconjunto de instancias de una causa.

En este contexto, tenemos dos conceptos a considerar:

**1.** La *consistencia*, que evalúa el grado en que las instancias que producen un resultado coinciden en mostrar la *condición causal* que se considera necesaria.
**2.** La cobertura, que evalúa la relevancia de la condición necesaria, es decir, el grado de qué instancias de la condición se combinan (*parean*) con instancias que producen el resultado. Esta discusión de las condiciones necesarias se basa en el trabajo de **Goertz (2002, 2003), Goertz y Starr (2002), y Braumoeller y Goertz (2000)**.

Estas evaluaciones del conjunto de datos establecidos, son importantes en el análisis de conexiones explícitas de la misma manera que las evaluaciones de significancia y la fuerza de la correlación lo son en el análisis de las conexiones correlacionales. Ambos conceptos: la *consistencia*, como la significancia, indican si una conexión empírica merece la atención del investigador. Si un hipotético subconjunto de datos no es consistente, entonces quiere decir, que la conjetura propuesta del marco teórico del investigador no la soporta. La cobertura, como la fuerza, indican la relevancia empírico o importancia de una conexión del marco teórico de la investigación. Como se mostrará en este capítulo, justo como es posible tener un correlación significativa pero débil, desde el punto de vista del análisis correlacional, también es posible, en el análisis del marco teórico de la investigación, tener un conjunto de datos establecido que sea altamente consistente pero de baja cobertura.

Estas mediciones propuestas al marco teórico de la investigación teorías de conjuntos, proporcionan herramientas vitales para refinar tanto al conjunto de datos nítidos como a los difusos, en un análisis de investigación social, con orientación a la administración de la innovación.

## Consistencia del marco teórico de la investigación

Las relaciones de conjuntos perfectamente consistentes son relativamente raras en la investigación social, con orientación a la administración de la innovación. Por lo general, pueden requerir datos pequeños de **N** o gran cantidad de datos o ambos. En general, los investigadores son capaces de identificar solo subconjuntos de datos debido a que las *excepciones*, están casi siempre presentes. Es importante, por lo tanto, desarrollar descripciones útiles de medidas del grado en que un conjunto de datos se ha aproximado,es decir, el grado en el cual, la evidencia es consistente con el argumento de que existe una conjunto de datos, como sigue:

1. Primero, se debe abordar la evaluación de la *consistencia* del conjunto de datos nítidos, donde simplemente una medición es suficiente y posteriormente se convierte en conjunto de *datos difusos*.
2. Al realizar evaluaciones de *consistencia,* es prudente tomar el número de casos en cuenta. La *consistencia* perfecta, no garantiza que exista una conexión al marco teórico de la investigación significativa. Por ejemplo, de la **Tabla 4.5** suponga, que tres de las doce empresas innovadoras clúster con calidad de votación por liderazgo tecnológico **W** no altas (casos: **C, I, J**, (**0.2**),(**0.2**) y (**0.0**)) producen niveles de influencia **A** no altos, pero no tanto como lo esperado (mismos casos**: C, I, J, (0.6), (0.6)** y **(0.8)**); la conclusión prudente, sería que esta conexión, aunque interesante y perfectamente coherente desde un punto de vista del marco teórico de la investigación, bien podría ser considerado una casualidad **(Ragin, 2000)**, por lo que es suficientemente justificado como para indicar, a un científico al menos, que puede existir algún tipo de conexión integral que es digno de mayor investigación.

3. El ejemplo mencionado, sugiere un análisis como medida directa de la *consistencia* del conjunto de datos nítidos, es decir, en la proporción de casos con un causa dada o combinación de causas que también producen el resultado. Con tres de tres casos consistentes, la proporción es **1.0**; con nueve de doce casos consistentes, la proporción es **0.75.** Como los explica **Ragin (2000)**, el **N** de casos puede ser tenido en cuenta mediante el uso de puntos de referencia y una probabilidad exacta de prueba. En general, los puntajes de *consistencia* deben ser tan cerca de **1.0** (*consistencia* perfecta) como sea posible. No se recomiendan puntajes de *consistencia* por debajo de **0.75**, ya que aún y cuando tengan soporte en su marco teórico de la investigación, que existe una relación establecida, son incluso, difíciles de comprobar **(Ragin, 2004)**.
4. La evaluación de la *consistencia* del conjunto de *datos difusos*, es más interesante y más desafiante que el conjunto de datos nítidos **(Ragin, 2000)** ya que, los casos pueden tener diversos grados de pertenencia en conjuntos de datos con puntajes que van de **0.0** a **1.0**.
5. Se dice que, existe una subconjunto de *datos difusos*, cuando los puntajes de pertenencia en un conjunto son consistentemente menores o iguales en otro. Por ejemplo, una empresa innovadora tienen un puntaje de pertenencia parcial dentro del clúster. La calibración del grado de pertenencia en el conjunto de *datos difusos,* involucra tanto una evaluación cuantitativa como cualitativa que debe ser referida a un marco teórico de la investigación, sustantivo en conocimiento **(Ragin, 2000)**. Como se explicó, un conjunto de *datos difusos*, existe cuando el puntaje de pertenencia en un conjunto es consistentemente menor o igual a sus correspondiente puntaje de pertenencia en otro.

**6.** Por ejemplo, retomando la **Tabla 1.2**

**Tabla 1.2. Caso hipotético: conexión basada en conexiones por correlación**

| Estatus | Falta plan de negocio | Cálculo erróneo de precio en producto/servicio |
|---|---|---|
| **Pyme innovadora menos exitosa (falla)** | 7 | 11 |
| **Pyme innovadora exitosa (sobrevive)** | 17 | 5 |

Fuente: Elaboración propia.

Si el grado de pertenencia en *falta plan de negocio* es consistentemente menor o igual al puntaje de pertenencia *Pyme innovadora* menos exitosa (falla), entonces la primera es un subconjunto de la segunda. Recordemos que con conjuntos de datos nítidos, no importa qué instancias de fracaso existan que no sean también, instancias de la falta de un plan de negocios, debido a que hay (hipotéticamente), otras formas de fallar (retomando la **Tabla 1.3**).

**Tabla 1.3. Caso hipotético: conexión basada en el marco teórico de la investigación**

| Estatus | Falta plan de negocio | Cálculo erróneo de precio en producto/servicio |
|---|---|---|
| **Pyme innovadora menos exitosa (falla)** | 15 | 16 |
| **Pyme innovadora exitosa, (sobrevive)** | 9 | 0 |

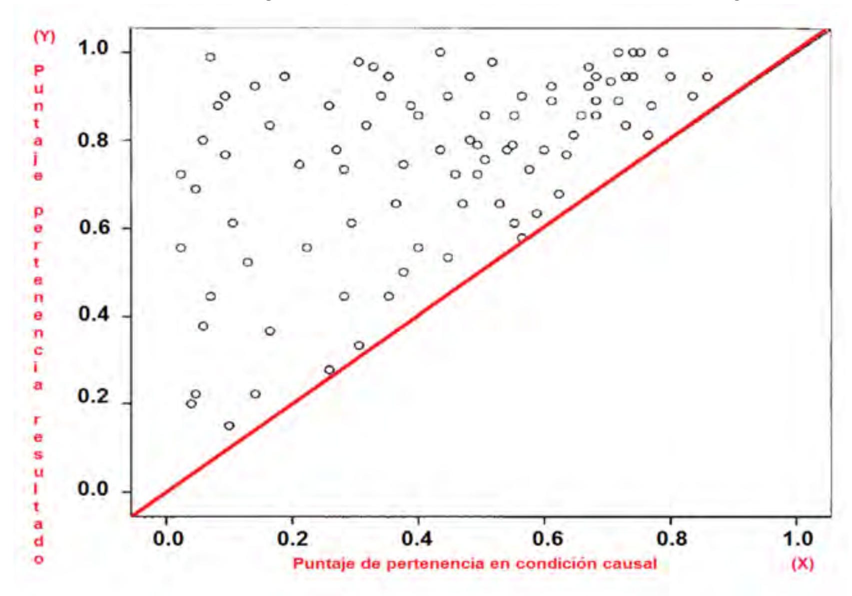
Fuente: Elaboración propia.

**7.** Con los conjuntos de *datos difusos* (*fuzzy-sets*), una situación paralela ocurre cuando los casos muestran resultados de puntajes de pertenencia, que exceden en gran medida sus puntajes en la *condición causal*. Por ejemplo, un caso pudiera tener una puntuación de **0.90** como pyme innovadora menos exitosa (falla) pero una puntuación de solo **0.20** en falta plan de negocios. A partir del análisis del conjunto de datos nítidos de

los argumentos del marco teórico de la investigación, este caso pudiera ser consistente porque puede haber varias formas de fallar, incluidas las rutas para pymes exitosas (sobreviven) con puntajes de pertenencia débil en el conjunto de las pymes.

8. Por el contrario otro caso que pudiera tener una puntuación de **0.30** como pyme innovadora menos exitosa (falla) pero una puntuación de **0.80** en falta plan de negocios, haría contradicción con los supuestos del marco teórico de la investigación.
9. El subconjunto de *datos difusos* (*fuzzy-sets*), tienen una forma triangular cuando se grafica en dos ejes difusos, como se muestra en la **Gráfica 5.1.**

**Gráfica 5.1. Subconjunto de datos difusos consistente en *es suficiente***




Fuente: Elaboración propia.

10. En esta figura, la causalidad en la condición **X** es un subconjunto del resultado **Y**; así, todos los valores $X_i$ son menores o igual que sus valores de $Y_i$ correspondientes, donde **i** indica referencia a valores

individuales de **X** o **Y** u observaciones específicas de **X** o **Y**. Tenga en cuenta que los casos en la esquina superior izquierda de la gráfica, no contradicen la idea de que esta causa pueda ser suficiente pero no necesaria para producir el resultado, porque estos casos tienen una alta pertenencia en el resultado debido a la operación de *condiciones causales* distintas de **X** (*un argumento de ser suficiente sin ser necesario*, permite múltiples trayectorias). Así, cuando el puntaje en **X** es baja, se permite una amplia gama de valores de $Y_i$**.** Cuando la pertenencia en **X** es alta, sin embargo, pueden existir muchas más oportunidades para violar al subconjunto de datos como rango permisible de valores $Y_1$ estrechos permitidos. Por supuesto, en un análisis correlacional convencional, los puntos en la esquina superior izquierda se considerarían errores y que a su vez socavaría la correlación entre **X** e **Y**.

**11.** Siguiendo a **Ragin (2000)**, la definición de la *consistencia* de un conjunto de *datos difusos* (*fuzzy-sets*) es claramente simple. En el gráfico, se observa el puntaje de pertenencia del resultado **Y** contra el puntaje de pertenencia en un *condición causal* o combinación de *condiciones causales* **X**, la *consistencia* se define como la proporción de casos en o por encima de la diagonal principal de la gráfica. Si la pertenencia en **X,** es consistentemente menor o igual que la pertenencia en **Y**, entonces todos los casos se graficarán en o sobre la diagonal principal, produciendo un puntaje de *consistencia* de **1.0** (o **100%** consistente). En el algoritmo de inclusión difusa (*fuzzy inclusion*) descrito en **Ragin (2000)**, la *consistencia* de los puntajes se calcula para diferentes combinaciones de *condiciones causales*, y estos puntajes proporcionan la base para evaluar el ser suficiente **(Ragin, et al., 2007; Ragin, 2007)**. Por ejemplo, si es significativamente mayor al 80% de los casos trazados en o sobre la diagonal principal en la gráfica que se acaba de describir, entonces el investigador podría afirmar que la causa o combinación causal **X** es casi siempre suficiente para producir el resultado **Y**.

**12.** Los procedimientos presentados por **Ragin (2003a)** para la evaluación del ser suficiente de las combinaciones causales se basan en el simple categorización de casos como consistente o inconsistente y el cálculo

de la proporción simple de casos consistentes. Un relativo problema se presenta, en los procedimientos que siguen de cerca la plantilla de conjunto nítido, procedimiento que se refiere al contraste entre casos con puntaje de pertenencia fuertes vs. débiles en la *condición causal* o la combinación de *condiciones causales* **X**. Específicamente, casos con puntaje de pertenencia fuerte y débil en la combinación causal se ponderan igualmente en el cálculo, sin embargo, difieren sustancialmente en su relevancia para la afirmación del marco teórico de la investigación y por lo tanto al argumento de que **X** es suficiente para **Y**. Por ejemplo, un caso con un puntaje de pertenencia de solo **0.25** en el conjunto de casos con la combinación causal **X** y una puntuación de **0.0** en el conjunto de resultados **Y** es solo tan inconsistente como un caso con una puntuación de **1.0** en la combinación causal y una puntuación de **0.75** en el resultado. Un puntaje de pertenencia de **0.25** indica que un caso está más afuera que dentro del conjunto; **0.5** es el punto de cruce. De hecho, sin embargo, el segundo caso inconsistente, con puntuación completa en **X**, claramente tiene más relación con el argumento del marco teórico de la investigación, porque es una instancia mucho mejor de la combinación causal. Esto constituye así, una in*consistencia* más evidente que el primer caso a pesar de la igualdad de brechas: los valores de $X_1$ exceden los valores de $Y_i$ en la misma cantidad.

**13.** El mismo razonamiento es válido para casos consistentes, de forma que:

   **a.** Un caso consistente con dos puntajes de pertenencia altos (por ejemplo, **0.9** en la combinación causal y **1.0** en el resultado) es claramente más relevante el argumento del marco teórico de la investigación que un caso consistente con dos puntajes bajos (por ejemplo, **0.1** en el combinación causal y **0.2** en el resultado) o

   **b.** Un caso es consistente con un puntaje bajo en la combinación causal (digamos, **0.15**) y puntaje alto en el resultado (digamos, **0.8**). Sin embargo, todos se cuentan por igual en la fórmula para la *consistencia* utilizada en **Ragin (2000)** (la proporción de casos en o por encima de la diagonal principal en la gráfica difusa).

Suponga que trata de apoyar una discusión en una presentación oral, usando evidencia de profundidad sobre un caso, con una pertenencia débil en los conjuntos relevantes. El sentido común, pensando que indica que esta presentación sería un desperdicio de tiempo, se propone formalizarla, precisamente con puntajes de pertenencia difusos. Casos con una fuerte pertenencia a la *condición causal*, proporcionan tanto los más relevantes casos consistentes como inconsistentes. Esta idea de sentido común, se operacionaliza en la medida alternativa de la *consistencia* de los datos de conjuntos de *datos difusos* (*fuzzy-sets*) con argumentos del marco teórico de la investigación, recomendados en este capítulo. Este procedimiento alternativo, como el propuesto en **Ragin (2000)**, hace la diferencia entre los casos consistente e inconsistente, usando la diagonal del gráfico, al plantear:

- Un caso en o encima de la diagonal principal, es consistente porque su pertenencia a la *condición causal* es menor o igual que su pertenencia en el resultado.
- Un caso debajo de la diagonal principal es inconsistente porque su pertenencia en la *condición causal* es mayor que su pertenencia en el resultado.

Sin embargo, en lugar de simplemente calcular la proporción de casos consistentes, el procedimiento alternativo utiliza puntajes de pertenencia difusos.

## Caso 1

Considere, por ejemplo, los datos hipotéticos de conjuntos de *datos difusos* (*fuzzy-sets*) sobre el grado de pertenencia en productos de innovación tecnológica y aplicaciones en el sector salud para doce empresas innovadoras mostradas.

**P1.** *¿Cuál es la consistencia de los datos mostrados en la* **Tabla 5.1***?*

**Tabla 5.1. Caso hipotético de un subconjunto de datos difusos (*fuzzy-sets*)**

<table>
<tr><th rowspan="2">Empresa innovadora</th><th>(X)<br>Productos de innovación tecnológica</th><th>(Y)<br>Resultado: Aplicaciones en el sector salud</th><th>(C)<br>Consistencia ( $X_i < = Y_i$ )</th></tr>
<tr><th>$\Sigma X_i = 4.7$</th><th>$\Sigma Y_i = 7.2$</th><th>$\Sigma [ \min ( X_i , Y_i ) ] / \Sigma (X_i) = 4.7 / 4.7$</th></tr>
<tr><td>1</td><td>0.7</td><td>0.9</td><td rowspan="12">En referencia a los puntos tratados en este apartado, se tiene:<br>• Punto 3:<br>Caso $X_i$ = 0.7, entonces (4.7 / 4.7)= 1.0<br>• Punto 4:<br>Caso $X_i$ = 1.0, entonces (4 / 5)= 0.8<br>• Punto 5:<br>Caso $X_i$ = 1.0, por lo que $Y_i$ se suma a $X_i$ produciendo: (4.9 / 5)= 0.98<br>• Punto 6:<br>Caso $X_i$ = 1.0; $Y_4$ = 0.3, por lo que $Y_4$ se suma a $X_i$ produciendo: (4.3 / 5)= 0.86</td></tr>
<tr><td>2</td><td>0.1</td><td>0.9</td></tr>
<tr><td>3</td><td>0.1</td><td>0.1</td></tr>
<tr><td>4</td><td>0.3</td><td>0.3</td></tr>
<tr><td>5</td><td>0.9</td><td>0.9</td></tr>
<tr><td>6</td><td>0.7</td><td>0.7</td></tr>
<tr><td>7</td><td>0.3</td><td>0.9</td></tr>
<tr><td>8</td><td>0.3</td><td>0.7</td></tr>
<tr><td>9</td><td>0.3</td><td>0.7</td></tr>
<tr><td>10</td><td>0.1</td><td>0.1</td></tr>
<tr><td>11</td><td>0.0</td><td>0.0</td></tr>
<tr><td>12</td><td>0.9</td><td>1.0</td></tr>
</table>

Fuente: Elaboración propia.

Para calcularlo, se asumen los siguientes supuestos:

1. Suponga que el puntaje para productos de innovación tecnológica en la primera fila de la **Tabla 5.1** fue **1.0** en lugar de **0.70**. Sería inconsistente con la relación establecida porque este valor excede el correspondiente la puntuación de pertenencia de resultado, 0.90. Mientras que, el conjunto de datos no debería tomar consistentemente por más tiempo los casos enumerados en la **Tabla 5.1**, aún sería muy cercano a ser perfectamente consistente, con once de los doce casos consistentes y solo una falta manifestada.
2. Una medida directa, de la *consistencia* del marco teórico de la investigación, usando puntajes de pertenencia difusa, es la suma de las *consistencias* en una *condición causal*, o combinación de *condiciones causales*, divididas por la suma de todos los puntajes de pertenencia en una causa o combinación causal **(Ragin, 2003)**.

3. En la **Tabla 5.1**, el valor de esta medida es **1.0 (4.7 / 4.7)** porque todos los puntajes de pertenencia en la *columna 1* son consistentes.
4. Si la puntuación para productos de innovación tecnológica en la primera fila de la **Tabla 5.1** es cambiado a **1.0**, la *consistencia* cae a **0.8 (4 / 5)**. El numerador es **1.0** como unidad difusa más baja que el denominador debido al puntaje inconsistente de **1.0**. La reducción de *consistencia* **a 0.8** (de *consistencia* perfecta, **1.0**) es sustancial porque **1.0** (el valor sustituido para **0.70** en la primera fila) es un puntaje de pertenencia grande. Esta medida de *consistencia* se puede refinar aún más para que proporcione crédito por errores y sanciones por puntajes de pertenencia causales que excedan sus marcas, las puntuaciones de pertenencia de resultado correspondientes por amplios márgenes. La fórmula descrita aquí es la implementada en el algoritmo de *tabla de verdad* difusa de **fsQCA** de **Ragin (et al., 2007)**. Este ajuste se puede lograr, agregando al numerador en la fórmula recién esbozada (la suma de la *consistencia* de los puntajes divididos por la suma de todos los puntajes de pertenencia en la causa o combinación causal) la parte de cada puntaje de pertenencia causal inconsistente que sea consistente con el resultado.
5. Por ejemplo, si la puntuación para productos de innovación tecnológica en la primera fila de la **Tabla 5.1** se cambian a **1.0**, posteriormente la mayor parte de su puntaje es consistente, hasta el valor del puntaje de pertenencia resultante **0.90**. Esta porción se agrega al numerador de la medida de *consistencia*. El uso de esta medida más refinada de *consistencia* produce un puntaje de *consistencia* general de **0.98 (4.9 / 5)**. Este puntaje de *consistencia* ajustada es más compatible con la evidencia. Después de todo, solo uno de los puntajes es inconsistente, y es una muy tendiente a falla. Por lo tanto, una *consistencia* se debe esperar una puntuación cercana a **1.0.** Tenga en cuenta, que la medida de *consistencia* revisada describe sustanciales sanciones para grandes *inconsistencias*.
6. Supongamos nuevamente, que la puntuación de productos de innovación tecnológica en la primera fila de la **Tabla 5.1** es **1.0**, pero esta vez suponga que el valor correspondiente del resultado es de aplicaciones en

el sector salud y es solo **0.3**. La porción consistente del puntaje de pertenencia **1.0** es **0.3**, produciendo una suma total de solo **0.3** al numerador. El puntaje de *consistencia* resultante en este caso sería **0.86 (4.3 / 5).** Este puntaje más bajo, refleja el hecho de que el puntaje inconsistente excede su objetivo por un amplio margen.

## Formalizando la consistencia

Es posible formalizar el cálculo de la *consistencia* teórica de conjuntos difusos como sigue:

$$\textbf{Consistencia} \ (X_i <= Y_i) = \Sigma \ [\min (X_i, Y_i)] / \Sigma (X_i)$$

Donde:
**min** indica la selección más baja de dos valores; $X_i$ es el grado de pertenencia en una combinación causal; $Y_i$ es el grado de pertenencia en el resultado **(Kosko, 1993; Smithson y Verkuilen, 2006)**. Cuando los valores $X_i$ son todos menores o iguales a sus valores de $Y_i$ correspondientes, el puntaje de *consistencia* es **1.00**; cuando solo unos pocos no cumplen la condición, el puntaje es ligeramente menor a **1.00**; cuando son muchos los que no la cumplen, hay puntuaciones inconsistentes con algunos valores de $X_i$ muy superiores a sus valores de $Y_i$ correspondientes, por lo que la *consistencia* puede caer por debajo de **0.5**. Tenga en cuenta que los mismos procedimientos para incorporar criterios probabilísticos, mencionados anteriormente y discutidos en detalle en **Ragin (2000)**, se pueden aplicar aquí. Estas pruebas probabilísticas, requieren un valor de referencia (por ejemplo, coherencia de **0.75**) y un nivel alfa (por ejemplo, significancia de **0.05**).

## Caso 2

**P2.** Muestre los gráficos correspondientes

1. Finalmente, cuando la fórmula para la *consistencia* del cálculo del marco teórico difuso de su investigación, se aplica como datos de conjunto de datos nítidos, devuelve la simple proporción de casos consistentes. Por lo tanto, la fórmula se puede aplicar a puntuaciones de pertenencia nítidas y difusas por igual. Esta misma fórmula general, también se puede aplicar a la evaluación de la *consistencia* de un conjunto de relaciones establecida que indica que una *condición causal* es una condición necesaria para un resultado.
2. Un argumento de ser necesario causal, es compatible cuando se puede demostrar que las instancias de un resultado constituyen un subconjunto de instancias de una *condición causal*. Con el conjunto de *datos difusos*, la *consistencia* de la relación de condición es necesaria depende del grado en que se puede demostrar que la pertenencia en el resultado es consistentemente menor o igual que la pertenencia en la causa, $Y_i <= X_i$. La **Gráfica 5.2,** ilustra este conjunto de *datos difusos*.

**Gráfica 5.2. Subconjunto de datos difusos consistente en *es necesario***

Fuente: Elaboración propia.

Donde el resultado **Y,** es un subconjunto de la *condición causal* **X**; por lo tanto, todos los valores de $Y_i$ son menores o iguales que sus valores $X_i$ correspondientes. Tenga en cuenta, que los casos en la esquina inferior derecha de la gráfica, no contradicen el ser necesario, ya que estos son casos que tienen baja pertenencia en el resultado porque carecen de alguna otra *condición causal* no especificada. Después de todo, lo *condición causal* en este ejemplo solo es necesaria, no suficiente.

3. Por supuesto, en un análisis correlacional convencional, los casos en la esquina inferior derecha se considerarían errores, lo que a su vez socavaría el correlación entre **X** e **Y**. Tenga en cuenta, sin embargo, que cuando la pertenencia en **X** es baja, la pertenencia en **Y** también debe ser baja. Por lo tanto, en el rango bajo de **X**, existen muchas oportunidades para violar el subconjunto de datos, con solo un rango estrecho de valores de $Y_i$ permitidos. Dada la desigualdad ($Y_i <= X_i$) que es lo contrario de la desigualdad que define el ser suficiente ($X_i <= Y_i$), una medida simple de la *consistencia* del subconjunto de datos para una *condición necesaria*, es:

$$\textbf{Consistencia}\ (Y_i <= X_i) = \Sigma\ [\ \min\ (X_i, Y_i)\ ]\ \Sigma\ /\ (Y_i)$$

Donde:
Cuando todos los valores de $Y_i$ son menores o iguales a sus valores de $X_i$ correspondientes, esta fórmula devuelve un valor de **1.0**. Cuando muchos $Y_i$ superan sus valores $X_i$ correspondientes por amplios márgenes, el cálculo devuelve un valor inferior a **0.5**. Por supuesto, es importante recordar que la interpretación de cualquier relación del marco teórico de la investigación, como *necesaria* o *suficiente* debe ser construido sobre una base sólida de conocimiento teórico y soporte empírico de soporte. Las conexiones causales no son inherentes a los datos. La *consistencia* del marco teórico de la investigación es solo una pieza de evidencia en el arreglo de apoyo que debe ser llevado a cabo, cuando un investigador hace una afirmación de ser suficiente o ser necesario o de cualquier otro tipo de conexión causal o constitutiva.

## Cobertura del marco teórico de la investigación

Cuando los investigadores permiten la equifinalidad **(George y Bennett, 2005; George, 1979)** y la *complejidad causal* **(Ragin,1987)**, un hallazgo común de ambas, es que un resultado dado puede ser producto de varias diferentes combinaciones de condiciones. Estas combinaciones, se entienden generalmente como trayectorias causales alternativas para producir un resultado. Generalmente, estas trayectorias alternativas se tratan como lógicamente equivalentes (es decir, como *sustituibles*). Sin embargo, es común en los análisis de conjuntos nítidos, evaluar la proporción de casos que siguen cada trayectoria, es decir, el número de casos siguiendo una trayectoria específica hacia el resultado dividido por el número total de instancias del resultado. Esta simple proporción, es una medida directa de cobertura del marco teórico de la investigación y es un indicador directo de la importancia empírica de una combinación causal. Claramente, una combinación causal que cubre o representa solo una pequeña proporción de instancias de un resultado no es tan empíricamente importante como uno que cubre una gran proporción. Tenga en cuenta, que la cobertura mide solo la importancia empírica, no la importancia teórica. Una relación del ser suficiente puede ser bastante rara desde un punto de vista empírico (y por lo tanto, exhiben una cobertura baja), y que aún todavía podría ser centralmente relevante para la teoría. Por ejemplo, la relación de ser suficiente podría ser prueba de que una trayectoria, que se creía empíricamente imposible, al menos desde la perspectiva de la teoría, de hecho no lo es.

La cobertura es distinta de la *consistencia*, y las dos a veces funcionan uno contra el otro porque la alta *consistencia* puede producir baja cobertura. Es decir, argumentos complejos a partir de los marcos teóricos de una investigación, que involucran la intersección de muchos conjuntos de datos, pueden lograr una *consistencia* notable pero con una cobertura baja. Por ejemplo, considere la formación académica de los CEO de las empresas innovadoras en clúster, expuestas en la **Tabla 4.5** que combinan excelentes registros académicos, puntajes de exámenes de alto rendimiento, padres con educación universitaria, altos ingresos de los padres, graduación de las

universidades de prestigio, etcétera. No sería sorprendente proponer, que el 100% de estas personas son capaces de evitar la pobreza con la *consistencia* del marco teórico de la investigación que parte la afirmación, es inusualmente perfecta basado a los datos a nivel individual, pero ciertamente, *no es imposible*. Sin embargo, hay relativamente pocos individuos, con esta combinación específica de circunstancias altamente favorables entre los muchos que exitosamente evitan la pobreza. Desde un punto de vista práctico, por lo tanto, este alto nivel de *consistencia* del marco teórico que sustenta la investigación, no es convincente, porque la combinación causal, está estrechamente formulada, de tal forma que resulta que su cobertura, es trivial.

Si bien, a menudo hay una compensación entre *consistencia* y *cobertura*, es razonable calcular la cobertura, solo después de establecer que un conjunto de relaciones es consistente. No tiene sentido calcular la cobertura de una causa o combinación de causas que no es un subconjunto consistente del resultado. Además, como quedará claro en la discusión que sigue, el mismo cálculo realizado al marco teórico de la investigación, tiene diferentes significados, dependiendo del contexto del cálculo. Por lo tanto, es importante cumplir con el protocolo descrito aquí, para que los resultados de las evaluaciones de *consistencia* y cobertura sean significativos: la *consistencia* del marco teórico de la investigación debe ser establecida, antes de evaluar la cobertura.

## Caso 3

**P3.** Realice análisis de *cobertura* con los datos de las **Tabla 5.2 y 5.3.**

**Tabla 5.2. Caso hipotético de la formación de un CEO de empresas innovadoras**

| Estatus del CEO de una empresa innovadora del clúster | Bajo logro educacional promedio | Alto logro educacional promedio | Criterios |
|---|---|---|---|
| **No en línea de pobreza económica** | **C1:** 3046 | **C2:** 1474 | • **Punto 1:** C2 / (C2 + C4) = 1474 / (1474+55) = 0.964 |
| **En línea de pobreza económica** | **C3:** 625 | **C4:** 55 | • **Punto 2:** C2 / (C1 + C2)= 1474/(3046+1474)= 0.0326 |

Fuente: Ragin (2008), con adaptación propia.

En dicha tabla, se muestra una hipotética idea de cobertura a partir de una tabulación cruzada del estado de pobreza económica (en línea vs. no en línea) contra el logro educativo promedio (alto vs. no alto), del CEO de una empresa innovadora del clúster, utilizando los conjuntos de datos nítidos y datos de nivel individual por CEO, porque:

1. Este análisis, utilizando datos binarios, apoya el argumento de que los CEOs con alto rendimiento educativo son capaces de evitar la pobreza. Esta afirmación del argumento que proviene de un elemental marco teórico, es apoyado por la alta proporción de casos en el segunda columna que no está en la pobreza. La celda **C2** dividida por la suma de las celdas **C2** y **C4**, que producen un puntaje de *consistencia* de: **1574/1529= 0.964**. Desde el punto de vista del marco téorico, la evidencia es de nuevo, altamente consistente.
2. Pero, ¿qué tan importante es esta trayectoria cuando se trata de evitar la pobreza económica? La forma más simple de contestar esta pregunta es calcular la proporción de los individuos que se encuentran en: no en línea de pobreza económica que tiene un alto logro educacional promedio, es decir, celda **C2** dividido por la suma de las celdas **C1** y **C2**, que es

**1474 / 4520= 0.326**. Este cálculo, muestra que la trayectoria en cuestión cubre casi un tercio de los casos no en línea de pobreza económica, que es sustancial.

Para fines de comparación, considere la **Tabla 5.3.**

**Tabla 5.3. Caso hipotético de la formación de un CEO de empresas innovadoras: frecuencias alteradas**

| Estatus del CEO de una empresa innovadora del clúster | Bajo logro educacional promedio | Alto logro educacional promedio | Criterios |
|---|---|---|---|
| **No en línea de pobreza económica** | **C1:** 4373 | **C2**: 147 | En referencia a los puntos tratados en este apartado, se tiene:<br>• **Punto 1:**<br>C2 / (C2 + C4) =<br>147 / (147+5) = 0.967<br>• **Punto 2:**<br>C2 / (C1 + C2)=<br>147/(4373+147)= 0.0325 |
| **En línea de pobreza económica** | **C3:** 675 | **C4:** 5 | |

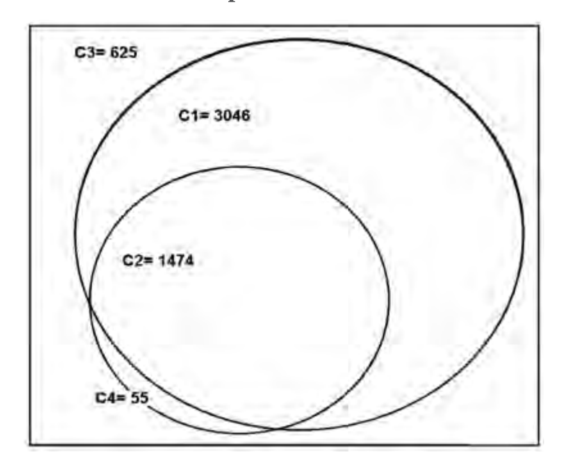
Fuente: Elaboración propia.

1. Dicha tabla, tiene el mismo número total de casos como en la **Tabla 5.2**, pero algunos de los casos han sido desplazados de la celda **C2** a la celda **C1** y de la celda **C4** a la celda **C3**. La proporción de casos consistentes con el argumento basado en el marco teórico de la investigación (**Tabla 4.3)** es **0.967**, casi lo mismo que en la **Tabla 5.2 (0.964).** Así, desde el punto de vista del marco teórico de la investigación, la evidencia es nuevamente muy consistente.
2. Pero ¿qué tan importante es esta trayectoria, utilizando las frecuencias hipotéticas presentadas en la **Tabla 5.3**? Su importancia se puede determinar calculando la proporción de casos que evitan la línea de pobreza económica que están cubiertos por argumentos de soporte basados en el marco teórico de la investigación, que es solo **0.0325 (147 / 4520)**. Así, en la **Tabla 5.3** el modelo del marco teórico de la investigación es nuevamente muy consistente, pero la cobertura es muy baja, lo que

indica (hipotéticamente) que tener un alto logro educacional promedio, no es una trayectoria importante hacia el resultado para evitar la línea de pobreza económica.

3. Los procedimientos para calcular la cobertura del marco teórico, usando el conjunto de *datos difusos*, establecen similitudes paralelas a los cálculos que se realizan al conjunto de datos nítidos, presentados anteriormente. Otra manera de entender el cálculo de la cobertura usando el conjunto de datos nítidos (celda **C2** dividida por la suma de las celdas **C1** y **C2**) es visualizar la **Tabla 5.2** como un diagrama de Venn que muestra una relación de subconjunto, como en la **Figura 5.1.**

**Figura 5.1. Diagramas de Venn que muestran el concepto de cobertura con el caso hipotético de la Tabla 5.2**




Notas:
Área **C1**. Casos con bajo logro educacional promedio, pero no en línea de pobreza económica.
Área **C2.** Casos con alto logro educacional promedio, pero no en línea de pobreza económica.
Área **C3.** Casos con bajo logro educacional promedio, pero en línea de pobreza económica.
Área **C4.** Casos con alto logro educacional promedio, pero en línea de pobreza económica.

Fuente: Ragin (2008) con adaptacion propia.

4. La idea básica detrás del cálculo de la cobertura, es evaluar el grado al cual, el subconjunto (el conjunto de casos con alto logro educacional promedio, en este ejemplo) cubre físicamente al conjunto más grande (el conjunto de casos no en línea de pobreza económica). Por lo tanto, la cobertura, el cual es un indicador de peso o importancia empírica, que puede verse como el tamaño del punto de cruce de los dos conjuntos en relación con el tamaño del conjunto más grande (que representa el resultado).
5. El cálculo del tamaño de punto de cruce de dos conjuntos de *datos difusos*, viene dado por su intersección:

$$\textbf{Punto de cruce} = \Sigma\ [\ \min\ (X_i\ ,\ Y_i\ )\ ]$$

   El cual, es lo mismo que el numerador en el cálculo de la *consistencia* del conjunto de *datos difusos* del marco teórico de la investigación, descrita anteriormente.
6. Con un conjunto de *datos difusos*, el tamaño del conjunto más grande (el denominador relevante) viene dado directamente por la suma de las puntuaciones de pertenencia en ese conjunto, es decir, la suma de las puntuaciones de pertenencia en el resultado:

$$\Sigma\ (Y_i\ )$$

   Este cálculo es similar y paralelo al simple conteo del número de casos en un conjunto (por ejemplo, el número de casos no en línea de pobreza económica) utilizando el conjunto de datos nítidos. Por lo tanto, la medida de la cobertura del conjunto de *datos difusos*, es simplemente la superposición expresada como una proporción de la suma de los puntajes de pertenencia, en el resultado:

$$\textbf{Cobertura} = (\ X_i < = Y_{\mathrm{i}}\ ) = \Sigma\ [\ \min\ (\ X_i\ ,\ Y_i\ )\ ]\ /\ \Sigma\ (Y_i)$$

7. En resumen, la fórmula para la cobertura de (**Y**) por (**X**):

$$\Sigma Y_i \textbf{ por } \Sigma X_i$$

Sustituye en el denominador de la fórmula para la *consistencia* de **Y** como un subconjunto de **X** (es decir, $Y_i <= X_i$ ), presentado en la discusión de la evaluación de la *consistencia* de un conjunto de datos de condiciones necesarias. Recuerde, sin embargo, que en el contexto actual (evaluación del ser suficiente), la cobertura de **Y** por **X** (es decir, como en la ecuación anterior) es calculado solo después de que se haya establecido que **X** es un subconjunto consistente de **Y**. Por lo tanto, el propósito del cálculo en el contexto del ser su iciente, es evaluar la magnitud de **X** en relación con **Y**, dado que la mayoría, si no todos, los valores de $X_i$ son menores o iguales que sus valores de $Y_i$ correspondientes.

**8.** En el contexto de las condiciones necesarias, por el contrario, el objetivo es evaluar la *consistencia* de **Y** como un subconjunto de **X**. Por lo tanto, en ese contexto, la expectativa es que la mayoría, si no todos, los valores de $Y_i$ serán menores o iguales a sus valores $X_i$ correspondientes, lo que indica una relación de condiciones necesarias. De hecho, si este no es el caso, entonces el resultado del cálculo tendrá un puntaje de *consistencia* (para $Y_i <= X_i$ ) que no llega a ser de *consistencia* perfecta (es decir, sustancialmente menor que **1.0**), indicando que **Y** no es un subconjunto consistente de **X**.

**9.** En resumen, el contexto debe tenerse en cuenta al realizar estas evaluaciones. Se deduce que es posible encontrar un:

$$\Sigma [ \min ( X_i , Y_i ) ] \text{ que está cerca de } \Sigma (Y_i )$$

produciendo una puntuación de cobertura muy alta, solo si los valores de $X_i$ son aproximadamente iguales a sus valores correspondientes de $Y_i$. Esta situación correspondería a una estrecha coincidencia de la dos conjuntos. Establecer la coincidencia no es lo mismo a correlación, sino que es un caso especial de correlación.

**10.** En una gráfica de dos conjuntos de *datos difusos*, cualquier línea recta que no sea vertical ni horizontal, produce un coeficiente de correlación perfecto. Sin embargo, solo ocurre una coincidencia perfecta cuando todos los casos se grafican exactamente en la diagonal principal de

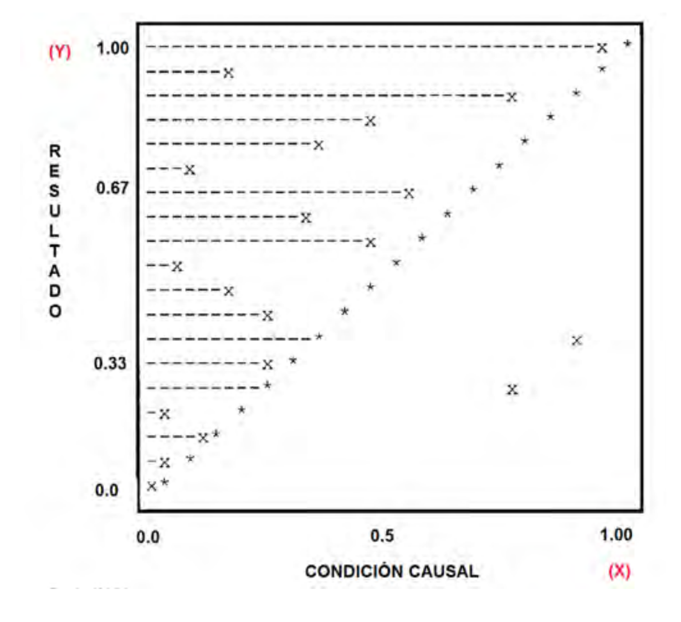
la gráfica difusa. Una medida simple del grado en que coinciden los puntajes de pertenencia en dos conjuntos es:

$$\Sigma\,[\,\mathbf{min}\,(\,X_i\,,\,Y_i\,)\,]\,/\,\Sigma\,[\,\mathbf{max}\,(\,X_i\,,\,Y_i\,)\,]$$

donde **max** indica el uso de la mayor de las dos puntuaciones. **Smithson y Verkuilen (2006)**, que contrastan la comorbilidad (*comorbidity*), la covarianza (*covariation*) y la co-ocurrencia (*co-occurrence*)).

La **Grá ica 5.3** muestra el concepto de *cobertura relevante* para el subconjunto de *datos difusos*, con $X_i <= Y_i$ **.**

**Gráfica 5.3. Ilustrando el concepto de cobertura**




Fuente: Ragin (2008) con adaptación propia.

**11.** Como en la **Gráfica 5.1**, la condición **X** es un subconjunto de resultado **Y**. Los puntos debajo de la diagonal principal constituyen violaciones de *consistencia* y, por lo tanto, socavan el argumento de que **X** es un subconjunto de **Y**. Sin embargo, solo hay dos puntos y el subconjunto

de relaciones, es en gran medida, consistente. Al calcular la cobertura, solo la parte del puntaje $X_i$ que está por encima de la diagonal principal, se cuenta como consistente (y por lo tanto, incluido como parte de la superposición entre **X** e **Y**). La mayoría de los puntos en la **Grá ica 5.3**, están por encima de la diagonal principal y, por lo tanto, son consis-tentes con $X_i$ **< =** $Y_i$. Cuando los valores de $X_i$ son pequeños en relación con sus valores correspondientes $Y_i$ están más cerca del eje **Y** que de la principal diagonal. Si bien estos puntos son consistentes con la relación de subconjunto $X_i$ **<=** $Y_i$, contribuyen relativamente poco a la cobertura, especialmente cuando los valores de $Y_i$**,** están por encima de **0.5**.

**12.** Las líneas horizontales punteadas en la **Gráfica 5.3**, muestran parte de los valores $X_i$ que cuentan como consistentes; estos valores se agregan al numerador de la fórmula para la cobertura. El denominador es la suma de los valores de $Y_i$. Las brechas de los valores consistentes de $X_i$ en la diagonal principal, muestran la parte del conjunto **Y** que no está cubierta por el conjunto **X**. El cálculo de la cobertura, también se puede aplicar a la evaluación de condiciones necesarias, donde el resultado es un subconjunto de la causa. **Goertz (2003, 2006)**, basándose en **Braumoeller y Goertz (2000)**, presenta un enfoque para la evaluación de las condiciones necesarias que aborda algunos de los mismos problemas discutidos en este capítulo. Una factor clave en su trabajo, es la distinción entre lo trivial y no trivial de las condiciones necesarias. Una condición trivial necesaria es aquella que está fuertemente presente en la mayoría de los casos, ya sea si estos casos o no lo muestran.

Por ejemplo, un nivel tolerable de quejas de los consumidores ante un producto/servicio, pueden ser una condición necesaria para que la organización realice la activación de acciones dentro de las empresas innovadoras, pero normalmente, un nivel tolerable de quejas casi siempre está presente en la operación, y la ausencia de quejas rara vez tiene la oportunidad de actuar como disparador de la empresa innovadora con acciones de innovación. Por lo tanto, se puede ver la existencia de un nivel tolerable de quejas de los consumidores ante un producto/servicio, como una condición nece-

saria empíricamente trivial. Por el contrario, nivel de liderazgo del CEO abierto y comunicativo (es decir, la ausencia de restricciones para actuar) podría verse como una condición no trivial necesaria, para que la organización realice la activación de acciones dentro de las empresas innovadoras, que normalmente se encuentran con restricciones para actuar. Mientras el cálculo específica recomendada en este capítulo, para evaluar la relevancia de las condiciones necesarias difiere en sus detalles por lo sugerido por **Goertz (2003, 2006)**, los objetivos esenciales permanecen similares. Los puntajes de pertenencia se dividen en el nivel de caso, y posteriormente estos radios, se promedian (*Ibídem*). En efecto, este procedimiento asigna casos de iguales importancia en el cálculo de una medida dada. En el enfoque expuesto aquí, sin embargo, los casos con puntajes de pertenencia bajos del conjunto de *datos difusos*, tienen menos peso porque son instancias débiles del fenómeno en cuestión. Esta estrategia de cálculo, hace que las medidas resultantes reflejen más los patrones observados, en los mejores casos.

Una medida simple de la importancia o relevancia de **X** como una condición necesaria para **Y** viene dada por el grado de *cobertura* de **X** por **Y**:

$$\textbf{Cobertura} = (\,Y_i <= X_i\,)\,\Sigma = [\,\min(\,X_i\,,\,Y_i\,)\,]\,/\,\Sigma\,(X_i)$$

Cuando la *cobertura* de **X** por **Y** es pequeña, el efecto restrictivo de **X** en **Y** es insignificante. Conceptualmente, una cobertura muy baja corresponde a una condición necesaria, empíricamente *irrelevante* o incluso sin sentido. Por ejemplo, casi todas las empresas innovadoras que recurren a apoyos de gobierno son del giro de alta tecnología, sería difícil afirmar que el giro de alta tecnología es la condición necesaria relevante (es decir, como estrictamente necesario de inicio) para el acceso a recurrir a apoyos de gobierno porque el conjunto del giro de alta tecnología es más pequeño que las empresas innovadoras. Por el contrario, cuando la cobertura de **X** por **Y** es sustancial, entonces el efecto restrictivo de **X** como condición necesaria, puede ser muy grande. Por ejemplo, si una proporción sustancial de las pymes quienes se asocian a empresas innovadoras, más tarde recurren a apoyos de gobierno y solo un muy pequeño número de pymes se vuelven empresas innovadoras, sin primero recurrir a apoyos de gobierno, entonces,

la cobertura es alta y la asociación de pymes con empresas innovadoras, puede considerarse una condición necesaria relevante, para ser una *pyme innovadora*. El contraste entre estas dos situaciones, de relevancia alta vs. baja en el análisis de las condiciones necesarias, se representa en la **Figura 5.2.**

**Figura 5.2. Diagrama de Venn ilustrando las condiciones necesarias**



Fuente: Ragin (2008) con adaptación propia.

La **Figura 5.2a,** representa una condición necesaria que ejerce alguna restricción en el resultado (la cobertura no es trivial). **La Figura 5.2b** representa empíricamente una condición trivial necesaria (muy baja cobertura por el marco teórico de la investigación). Usando el conjunto de *datos difusos*, la situación representada en la **Figura 5.2b** aparecería como una gráfica en la que, casi todos los casos tienen una pertenencia muy fuerte en **(X)** (la *condición causal*) y, por lo tanto, se trace en el extremo derecho **(Goertz 2003, 2006)**.

Al igual que la evaluación de la cobertura de una condición de ser suficiente, es importante evaluar la relevancia de una condición del ser necesario (es decir, su impacto de restricción) solo después de establecer, que el subconjunto de datos, es consistente. Es decir, primero debe establecerse que **Y** es un subconjunto consistente de **X** antes de evaluar el tamaño de **Y** en relación con el tamaño de **X.** El seguir este protocolo, evita la confusión, con respecto a la interpretación de lo que son cálculos esencialmente idénticos:

1. El cálculo de la *consistencia* de una relación del ser suficiente, es idéntica al cálculo de la cobertura (*relevancia*) de una relación de ser necesario, mientras que,
2. El cálculo de la cobertura de una relación de ser suficiente es idéntica al cálculo de la *consistencia* de una relación de ser necesario. El protocolo para evaluar la *consistencia* y cobertura para estos dos tipos de relaciones de contenidos del marco teórico de la investigación. Vea **Tabla 5.4.**

**Tabla 5.4. Protocolo de evaluación de la consistencia y cobertura**

| Tipo de relación del marco teórico de la investigación | | |
|---|---|---|
| **Procedimiento** | **La causa X es un subconjunto de resultados Y (ser suficiente)** | **El resultado Y es un subconjunto de causa X (ser necesario)** |
| **Etapa 1** | Evalúe la *consistencia*, con: $\Sigma[\min(X_i, Y_i)] / \Sigma(X_i)$ | Evalúe la *consistencia*, con: $\Sigma[\min(X_i, Y_i)] / \Sigma(Y_i)$ |
| **Etapa 2** | Si es consistente, evalúe la *cobertura*, con: $\Sigma[\min(X_i, Y_i)] / \Sigma(Y_i)$ | Si es consistente, evalúe la *cobertura*, con: $\Sigma[\min(X_i, Y_i)] / \Sigma(X_i)$ |

Fuente: Ragin (2008) con adaptación propia.

## Cobertura de partición

Cuando más de una condición o combinación de condiciones es suficiente para un resultado (es decir, cuando hay *equifinalidad*), la evaluación de la cobertura de combinaciones causales alternativas, proporciona evidencia directa de su relativa importancia empírica. Además, las evaluaciones de cobertura totales (*bruta*) se pueden complementar con evaluaciones de cada *cobertura única* de la combinación, ya que es posible dividir o particionar la cobertura del análisis del marco teórico de la investigación, de manera análoga a la partición de variación explicada en la regresión múltiple. La discusión de la partición de la cobertura que se presenta aquí, supone que el investigador ha demostrado que las *condiciones relevantes o combinaciones de las condiciones*, son subconjuntos consistentes del resultado.

## Caso 4

**P4.** ¿Cuáles la explicación a nivel de conjunto de *datos difusos*?

Por ejemplo, considere la evidencia del análisis de datos del conjunto de *datos difusos*, del caso hipotético mostrado en la **Tabla 5.5.**

**Tabla 5.5. Caso hipotético formación de jóvenes CEO innovadores: cálculo de cobertura**

| *Condiciones causales* | X Suma de puntuaciones de *consistencia* | Y Suma de puntuaciones de resultado | Cobertura (%) ( X/Y *100 ) |
|---|---|---|---|
| **T * I** | 307.387 | 1385.25 | 22.19 |
| **C** | 548.559 | 1385.25 | 39.60 |
| **T * I + C** | 607.709 | 1385.25 | 43.87 |

Fuente: Elaboración propia.

Se trata de un reporte de jóvenes en formación como CEO innovadores que se realiza por parte de una cámara de la industria electrónica y del software en un país emergente, el cual se realiza con:

**1.** El resultado, es el conjunto de *datos difusos* de aquellos con emprendimientos innovadores que tienen ingresos, asegurando su operación por más de dos años (**~ P**, indica el grado de pertenencia en el conjunto de datos de emprendimientos exitosos de dos años y **~** indica negación). Las tres *condiciones causales* parten del conocimiento del conjunto de *datos difusos* de casos con puntajes altos, en emprendimientos innovadores exitosos de más de dos años, por jóvenes CEO en formación, con:

   **a.** Un modelo de negocios definido, **T.**

   **b.** Con reporte de altos ingresos **I**, y

   **c.** Con educación universitaria **C,** de los jóvenes CEO.

**2.** Aplicando el **fsQCA (Ragin,et al., 2007)** a estos datos, se obtienen dos trayectorias, para evitar el bajo rendimiento de un emprendimiento innovador para al menos dos años de operaciones, a saber,

   **a.** La combinación de puntajes altos en la combinación del modelo de negocios y el reporte de altos ingresos de las empresas innovadoras en dos años, esto es: **( T * I )**, y

   **b.** La educación universitaria **(C),** por sí misma.

**3.** El cálculo de la cobertura total (bruta) de estas dos trayectorias para el resultado, de emprendimientos exitosos asegurando su operación por más de dos años **(~ P),** se muestra en la **Tabla 5.5**. La primera fila, reporta el cálculo de la cobertura para la combinación de puntuaciones altas en la combinación del modelo de negocios y el reporte de altos ingresos de las empresas innovadoras en dos años, esto es: **(T * I )**. La suma del punto de cruce o superposición entre **( T * I )** y el resultado es **307.387**. Las suma de las pertenencias en el resultado es 1385.25. Por lo tanto, esta combinación cubre aproximadamente el 22.19% del total de las pertenencias en el resultado **(307.387 / 1.385,25 = 0.2219).** Usando estos mismos procedimientos, la condición **C** cubre aproximadamente el 39.6% del total de la pertenencia del resultado (ver la fila dos de la **Tabla 5.5**). Por lo tanto, ambas combinaciones cubren una proporción

sustancial del resultado. Sin embargo, la cobertura total (bruta) de la condición **C** (educación universitaria) es mucho mayor.

4. Para fines de comparación, la **Tabla 5.5** también muestra la cobertura de dos combinaciones **( T * I y C )** concebidas como trayectorias alternativos hacia el mismo resultado, utilizando al operador lógica **OR**. Cuando las combinaciones causales se unen por un operador lógica **OR**, el puntaje de cada caso en la unión es el valor máximo de las dos trayectorias (es decir, el mayor de los dos puntajes, pertenencia en **( T * I )** vs. Pertenencia en **C**). En otras palabras, cuando hay más de una trayectoria a un resultado, es posible calcular qué tan cerca está un caso del resultado al encontrar su puntaje de pertenencia más alto entre las posibles trayectorias.
5. El grado de *cobertura* del resultado, por este máximo puntaje, a su vez, puede calcularse utilizando los mismos procedimientos aplicados, por separado, a los dos componentes. Este cálculo se muestra en el tercer fila de la **Tabla 5.5**, que reporta una cobertura de **43.87%,** mayor que la *cobertura* de cualquiera de los dos componentes (compare la fila 3 de **Tabla 5.5** con las dos primeras filas). Sin embargo, la cobertura de las dos trayectorias del modelo **(43.87%)** es moderadamente superior a la *cobertura total* (bruta) de la mejor trayectoria única (ruta **C**, con **39.6%**).
6. La **Tabla 5.5,** proporciona toda la información necesaria para la cobertura de partición, siguiendo el procedimiento que se utiliza en la partición de la varianza explicada del análisis de regresión múltiple. Para evaluar una contribución, separada o única de la variable independiente en una regresión múltiple que involucra varias variables predictoras correlacionadas, los investigadores calculan la disminución en la varianza explicada que ocurre una vez que la variable en cuestión, se elimina completamente de la ecuación multivariante especificada **(Mejía-Trejo, 2019**. Capítulo 5). Por ejemplo, para encontrar la contribución única de $\mathbf{X_i}$ en la varianza explicada en **Y,** es necesario calcular la ecuación de regresión múltiple con todas las variables independientes relevantes incluidas y posteriormente, recalcular la ecuación excluyendo $\mathbf{X_i}$.

La diferencia en la varianza explicada, entre estas dos ecuaciones, mostrará la contribución única de $X_i$. Estos procedimientos aseguran *que la varianza explicada que* (comparte con variables independientes correlacionadas, no se debe a $X_i$..

7. El objetivo de la partición en análisis de conjuntos de *datos difusos*, por el contrario, es evaluar la importancia relativa de las diferentes combinaciones de *condiciones causales relevantes*. Por lo tanto, el problema en los análisis de los marcos teóricos de investigación, no son las variables independientes correlacionadas, porque las *condiciones causales* no son analizadas de manera aislada de una variable con otra, ya que están involucradas en un análisis de regresión múltiple. Más bien, la cobertura de partición es importante porque algunos casos se ajustan a más de una trayectoria. Usando nuestro ejemplo, hay seguramente muchos investigadores quienes utilicen, altos puntajes en la combinación del modelo de negocios **T,** el reporte de altos ingresos de las empresas innovadoras en dos años **I** y la educación universitaria de los jóvenes CEO en formación **C.**

## Caso 5

**P5.** ¿Cuáles la explicación a nivel de conjunto de datos nítidos?

Supongamos, que un investigador encuentra que dos combinaciones de condiciones generan el resultado:

**Y : ( A * B ) AND ( C * D )**

Entonces:

1. El investigador calcula la cobertura de estas dos trayectorias y encuentra que el primero, abarca el **25%** de las instancias de **Y** (*cobertura* = **0.25**), mientras que el segundo, abarca el **30%** (*cobertura* = **0.3**). Sin embargo, al calcular su cobertura como trayectorias alternativas (es decir, su unión: **(A * B + C * D)**, donde la suma indica lógica **OR**, el investigador

encuentra que juntos abarcan solo el 35% de las instancias del resultado (cobertura = **0.35**).

2. La razón por la que esta cantidad es sustancialmente menor que la suma de los dos puntajes de cobertura separados (es decir, **0.35 < 0.25 + 0.3 = 0.55**) es porque las dos trayectorias se superponen parcialmente. Es decir, hay casos que combinan las cuatro *condiciones causales* (es decir, instancias de **A * B * C * D**) y el la cobertura de estas instancias se cuenta dos veces cuando se calcula la cobertura total (bruta) por separado, para las dos combinaciones causales. Afortunadamente, es cuestión muy simple de dividir la cobertura total (0.35 en este ejemplo) dentro de sus tres componentes:
   **a.** Únicamente debido a **( A * B ),**
   **b.** Únicamente debido a **( C * D )** y
   **c.** La superposición (es decir, debido a la existencia de casos de **(A * B * C * D)**.
3. Se puede calcular la *cobertura única* de cada término, por sustracción, siguiendo la plantilla proporcionada por el análisis de regresión; así:
   **a.** La *cobertura única* de la trayectoria **( A * B ),** es **0.35 - 0.3 = 0.05**;
   **b.** La *cobertura única* de la trayectoria **( C * D ),** es **0.35 - 0.25 = 0.10**; y
   **c.** El resto la cobertura total, se debe a la superposición entre estos dos términos.
4. En resumen, estos simples cálculos indican que el 20% de las instancias del resultado, son **( A * B * C * D );** el **5%** de las instancias del resultado son **( A * B )** sin **( C * D )** ; y el **10%** son **( C * D )** sin **( A * B ).**
5. El cálculo de la *cobertura única* de una combinación de condiciones. en el análisis de conjuntos de dato difusos, es similar o paralelo, a como se muestra en la **Tabla 5.6**, que utiliza el mismo nivel de datos individuales, utilizados en la **Tabla 5.5.**

**Tabla 5.6. Caso hipotético formación de jóvenes CEO innovadores: cálculo de partición de *cobertura***

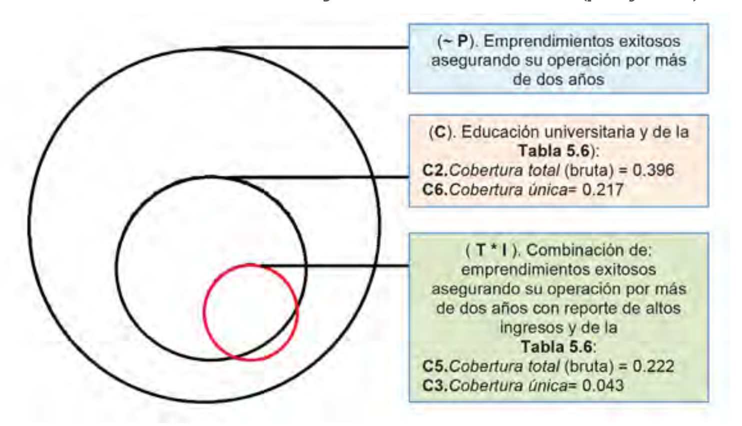
| Estatus | (X) Cobertura total | (Y) Sin términos | (X - Y) Única | Criterio |
|---|---|---|---|---|
| **Única a (T * I )** | **C1:** 0.4387 | **C2:** 0.3960 | **C3:** 0.0427 | En referencia a los puntos tratados en este apartado, se tiene:<br>• **Punto 10:**<br>C1 - ( C3 + C6 )= 0.1792 |
| **Única a ( C )** | **C4:** 0.4387 | **C5:** 0.2219 | **C6:** 0.2168 | |

Fuente: Elaboración propia.

6. La cobertura del resultado de emprendimientos exitosos asegurando su operación por más de dos años **(~ P)** que se debe únicamente a la ruta **( T * I)** es la diferencia entre la cobertura de las dos trayectorias del modelo **(0.4387)** y la cobertura que se obtiene una vez que **( T * I )** se elimina del modelo de dos trayectorias. Cabe destacar que en este ejemplo, es equivalente a la *cobertura* del otra trayectoria **C** por sí mismo.
7. Por lo tanto, la *cobertura única* de la trayectoria **( T * I )** es **0.0427**, es decir, **0.4387** (la cobertura combinada de las dos trayectorias) menos **0.3960** (la *cobertura única* de la trayectoria **C**).
8. Asimismo, la cobertura del resultado que se debe únicamente a la trayectoria C, es la diferencia entre la cobertura del modelo de dos trayectorias **0.4387** y la cobertura de la trayectoria **( T * I )** por sí misma **(0.2219),** o **0.2168**.
9. Estos cálculos revelan que la a *cobertura única* de la ruta C es mucho mayor que la *cobertura única* de la trayectoria **( T * I ).** De hecho, la cobertura de **( T * I )** es casi completamente, un subconjunto de la *cobertura* de C. (En otras palabras, la mayor parte de **T * I** es **T * I * C**).
10. Gran parte de la cobertura del modelo de dos trayectorias se superpone o traslapa. Esta proporción, se puede calcular con la diferencia entre la cobertura del modelo de dos trayectorias **(0.4387)** y la suma de las dos porciones únicas **(0.0427 + 0.2168 = 0.2595)**, que es **0.1792.**

**11.** La **Figura 5.3** ilustra estos resultados usando un diagrama de Venn. Cuando varias trayectorias diferentes pueden conducir al mismo resultado, es importante calcular tanto la cobertura total (bruta) y única de cada combinación causal. Estos cálculos a menudo revelan que solo existen unas pocas coberturas de combinaciones causales de nivel alto, incluso en análisis que tienen varias combinaciones del ser suficiente. Si bien es útil conocer todas las diferentes combinaciones causales vinculadas a un resultado, también es importante evaluar su peso empírico relativo. Los cálculos de *cobertura total* (bruto) y única proporcionan directamente estas evaluaciones.

**Figura 5.3. Caso hipotético formación de jóvenes CEO innovadores: diagrama de Venn de la partición de la cobertura del marco teórico usando el conjunto de datos difusos (*fuzzy-sets*)**




Fuente: Ragin (2008) con adaptación propia.

Estas medidas de la *consistencia* y la *cobertura* del marco teórico tienen muchos usos principalmente en la construcción de *tablas de verdad* nítidas, a partir de *datos difusos*, que es la base del algoritmo de *tabla de verdad* difusa.

## CAPÍTULO 6.

# La calibración y su importancia en fsQCA

El análisis por conjunto de datos difusos, es relativamente nuevo en el ámbito de las ciencias sociales y particularmente en la administración de la innovación. La primera obra registrada que ofreció una introducción de aplicaciones en la ciencias socialesl del mismo, es el trabajo de **Smithson (1987)**. Sin embargo, las aplicaciones ffueron escasas y distantes hastaeque los principios básicos del análisis de conjuntos de sdatos difusos, se elaboraron a través del análisis cualitativo comparativo (**QCA**) (**Ragin, 1987, 2000**), esto es, un análisis sistemático que se establece particularmentelen el marco teórico de una investigación, en oposición al análisis correlacional,tanto en su fundamentación como en su diseño. La asociación de estos dos, el **QCA** y el análisis del conjunto de datos difusos, produjo el análisis cualitativo comparativo por conjuntos de datos difusos (**fsQCA**.*fuzzy-sets qualitative comparative analysis*). Es una familia de métodos que ofrece una alternativa a los métodos cuantitativos convencionales, que se basan casi exclusivamente en razonamientos correlacionales.

La clave que hace muy útil el análisis de relaciones de conjuntos difusos, es la construcción fundamentada de modelos bien construídos a partir de su *calibración.* Pero, ¿cómo el investigador debe calibrar el grado de pertenencia, por ejemplo, en el conjunto de CEOs que presentan liderazgo innovador?, ¿cómo debe definirse este conjunto?, ¿qué constituye la plena pertenencia a un conjunto de datos?, ¿Qué constituye la no pertenencia completa?, ¿qué sería de un CEO con, digamos, un **0.75** de pertenencia en este conjunto (por ejemplo, más adentro que afuera, pero no totalmente dentro)?, ¿cómo diferiría este CEO de alguien con **0.9** de pertenencia? El mensaje principal de este capítulo, es que los conjuntos de *datos difusos*, a diferencia de las variables convencionales, deben calibrarse. Dado que deben ser calibrados, son superiores en muchos aspectos a los medidas convencionales, como se usan hoy en día, tanto en los enfoques cuantita-

tivo como en cualitativo en las ciencias sociales y particularmente, en la orientación de la administración de la innovación. En esencia, se sostiene que los conjuntos de *datos difusos*, ofrecen un forma intermedia entre la medición con enfoque cuantitativo y cualitativo. Sin embargo, esta forma de análisis intermedio, no es un compromiso entre estos dos enfoques; más bien, trasciende muchas de las limitaciones de ambos.

## Implicaciones de la calibración

La calibración, es una práctica necesaria y rutinaria en la investigación de las ciencias *duras* basadas en la ingeniería como la química o la física **(Pawson 1989, pp. 135–37)**. En estos y otras ciencias naturales, los investigadores calibran sus dispositivos de medición y las lecturas que producen estos instrumentos, ajustándolos para que coincidan o se ajusten a estándares confiables conocidos. Estos estándares hacen que las mediciones sean directamente interpretables **(Byrne, 2002)**. Por ejemplo, una temperatura de **26** grados Celsius es interpretable porque está situada entre **0** grados (el agua se congela) y 100 grados (el agua hierve). Sin embargo, la calibración de las medidas, de acuerdo con estándares acordados, es relativamente raro en las ciencias sociales especialmente en lo referente a la administración de la innovación. Existen varios esfuerzos por hacer calibraciones en las ciencias sociales reflejados en diversos reportes que van, desde la medición de la pobreza por ejemplo, con el Human Development Index **(HDI, 2018)** de Naciones Unidas, hasta el de la medición del grado de innovación por países reflejado en el *Global Innovation Index* de INSEAD-WIPO **(GII, 2019)**. En economía, por el contrario, la *calibración* tiene un significado diferente en general ya que los investigadores *calibran* los parámetros en los modelos, fijándolos en valores particulares, para que las propiedades y el comportamiento de otros parámetros en el modelo, puedan ser observados. Este tipo de calibración, es muy diferente al de la calibración explícita de medidas, preocupación central de este capítulo.

Normalmente, la mayoría de los científicos sociales están satisfechos con usar medidas no calibradas, y por lo tanto, simplemente muestran las posiciones de casos relativos entre sí. Sin embargo, en principio, una medidas no calibradas es claramente inferior a las medidas calibradas. Con una medida no calibrada de temperatura, por ejemplo, es posible saber que un objeto tiene una temperatura más alta que otra o incluso que tiene una temperatura más alta que el promedio para un conjunto dado de objetos, pero aún no se sabe si hace calor o frío. Del mismo modo, con una medida de innovación sin calibrar, es posible saber que una organización practica actividades de innovación más que otra, o más que el promedio, pero aún no se sabe si es una empresa innovadora o sigue siendo una empresa tradicional, con prácticas esporádicas de innovación. La calibración es especialmente importante, en situaciones donde una condición establece o da forma al contexto para otras condiciones. Por ejemplo, la relación entre la temperatura y el volumen en los cambios del agua, cualitativamente a **0** grados Celsius y posteriormente, nuevamente a **100** grados Celsius. El volumen disminuye, a medida que la temperatura tiende **0** grados Celsius y posteriormente aumenta, a medida que la temperatura tiende a 100 grados Celsius. La escala Celsius, es a propósito diseñada, como base de calibración para indicar estos cambios de fase, y los investigadores que toman en referencia dicha escala, saben de tomar en cuenta estos dos puntos de corte cualitativos, como referencia en sus trabajos de investigación.

## Condiciones de alcance y contexto

Las condiciones de contexto que operan como cambios de fase abundan en el estudio de los fenómenos sociales. La condición de establecimiento de contexto más básica, es la condición del alcance **(Walker y Cohen, 1985)**. Cuando los investigadores declaran que cierta propiedad o relación, tiene o existe solamente para casos de cierto tipo (por ejemplo, solo las personas con gran talento, trabajan para empresas innovadoras), han utilizado una condición de alcance para definir un contexto habilitador. Otro ejemplo

de una condición de contexto en las ciencias sociales con orientación a la admnistración de la innovación, es la consideración de los ecosistemas de innovación empíricas como condiciones propicias. Por ejemplo, cuando los investigadores argumentan que una propiedad o relación es válida solo para ecosistemas de innovación regionales, han utilizado una propuesta empírica de población como condición de contexto. Mientras que la distinción, entre las condiciones de alcance y las poblaciones a veces son difusas, su uso como condiciones de contexto suelen ser similares, paralelas. En ambos usos, actúan como condiciones que habilitan o deshabilitan propiedades o relaciones específicas. Las pruebas de interacción estadística, generalmente están motivadas la preocupación de las condiciones que alteran las relaciones entre otras variables, es decir, por las condiciones de establecimiento de contexto. Si el el efecto de **X** sobre **Y** aumenta de, ningún efecto a un efecto sustancial a medida que el nivel de una tercera variable **Z** aumenta, posteriormente **Z** opera como una condición de contexto, permitiendo la relación entre **X** e **Y.** A diferencia de las condiciones del alcance y límites de población, la variable de interacción **Z** en este ejemplo, varía según el nivel y no precisamente por una simple dicotomía de presencia / ausencia. De hecho, se podría argumentar que las condiciones dicotómicas de establecimiento de contexto tales como las condiciones del alcance son casos especiales de interacción estadística.

El hecho de que la variable de interacción **Z** varía según el nivel, como condición de contexto, plantea automáticamente el problema de la calibración, por lo que se crean preguntas como:

**1.** ¿En qué nivel de **Z** se hace posible una relación entre **X)** e **Y**?,
**2.** ¿En qué nivel de **Z** existe una conexión fuerte entre **X** e **Y**?

Para responderlas, es necesario especificar los valores relevantes de **Z,** que es hablar de una calibración de facto de **Z**. En un rango específico de valores de **Z,** no hay relación entre **X** e **Y**, mientras que en otro rango existe una fuerte relación entre **X** e **Y**. Quizás sobre intermedio valores de **Z**, hay una relación débil a moderada entre **X** e **Y**. Para especificar estos valores o

niveles, es necesario incorporar, conocimiento externo sustantivo de alguna manera, para interpretar estos diferentes niveles como condiciones de contexto. Los investigadores que se aplican a la interacción estadística, por lo general, ignoran en gran medida, este problema y se han conformado con realizar pruebas amplias de interacción estadística, centrándose en incrementos a la varianza explicada de una variable dependiente, sin atender los problemas de calibración y contexto **(Ragin, 2008)**. Para preparar el escenario para la discusión de conjuntos de *datos difusos* y su calibración, se sugiere:

**1.** Primero, examinar las prácticas comunes de medición, tanto de los enfoques cuantitativo y cualitativo.
**2.** Después de examinar estas prácticas, deben mostrarse qué conjuntos de *datos difusos,*
**3.** Coinciden tanto con:
   **a.** El punto de vista de la calidad de la investigación cualitativa, cuyo objetivo es, distinguir entre variación *relevante e irrelevante* (es decir, interpretar la variación), y
   **b.** El punto de vista de la calidad de la investigación cuantitativa: cuyo objetivo es, la colocación precisa de los casos relacionados entre sí.

## La investigación cuantitativa: prácticas de mediciones comunes

La medición, como se está decribiendo hasta este momento y tal como se practica hoy en las ciencias sociales y en particular, en la administración de la innovación, ha permanecido con descuido y aplicación de falta de rigor sistemático, a pesar de que a fines del siglo pasado, se ha insistido en su adopción **(Duncan, 1984; Pawson, 1989)**. El enfoque dominante es el de modelo teórico basado en indicadores, en el cual, los científicos buscan identificar los mejores modelos basados en indicadores empíricos posibles, de sus conceptos teóricos.

Por ejemplo, **Dziallas y Blind (2019)** detectaron, tan solo con la palabras clave: *indicator, factor, determinant, innovation* aplicadas a las tres bases de datos principales de conocimiento **Science Direct, Web of Science y Scopus,** detectaron **1796** artículos potenciales, los cuales fueron validados y filtrados obteniendo **226** artículos que sirvieron de base a su modelo teórico basado en indicadores para finalmente proponer y bajo su punto de vista, una nueva categoría de indicadores de innovación de procesos. La gran mayoría de los trabajos analizados, se basan en indicadores provenientes de escala de intervalo (escalas de **Likert** principalmente) o de razón. Particularmente la *escala de razón,* se considera la forma más alta de medición, ya que está referenciado a un punto de cero absoluto (por ejemplo, **0 Km/Hr**); es importante observar, que un conjunto de *datos difusos*, tiene al menos tres valores de referencia: **1.0** (*completa pertenencia*), **0.0** (*completa no pertenencia*), **0.5** (el *punto de traslape o superposición* de casos que están más dentro vs. más fuera, del conjunto de datos en cuestión). Si se acepta que tales referencias, señalan un mayor nivel de medición, posteriormente se deduce que un conjunto de *datos difusos*, es de mayor nivel de medición que una de escala de razón.

Las *escalas de intervalo* (*cero arbitrario*) y/o de *razón* (*con cero absoluto*), son considerados como *pertinentes y adecuados* para la aplicación de técnicas analíticas que evalúan las relaciones entre factores, tales como el análisis de regresión múltiple y las técnicas lineales relacionadas. Cabe señalar, que las versiones más sofisticadas del enfoque por modelo teórico basado en indicadores, están en los trabajos previos de la teoría psicométrica de **Nunnally y Bernstein (1994)**, el cual considera el concepto de variables o factores subyacentes y que se ha llevado a múltiples campos como el de la innovación. La idea central, es que un índice (uno o varios factores dependientes) al estar compuesto de múltiples indicadores correlacionados del mismo concepto subyacente, es altamente probable que sea más preciso y confiable, que cualquier indicador individual.

Sin embargo, el uso de múltiples indicadores tiende a sobrecalificar, al índice o factor subyacente de lo que realmente es **(Ragin, 2008)**, por lo que se consideran anomalías. Tales anomalías desafían la validez aparente

(*face-validity*) **(Mejía-Trejo, 2019b)** de un modelo teórico basado en indicadores, ya que se describirían sujetos que tienen puntajes (cargas factoriales) relativamente más altos, concluyendo niveles por ejemplo, de desarrollo de innovación que en realidad, no lo son. Idealmente, los diversos indicadores de un concepto, índice o factor subyacente, deben correlacionarse fuertemente entre sí. De hecho, la **OECD (2008)** ha hecho mandatorio el seguir un diseño de construcción de indicadores múltiples. Ver **Tabla 6.1.**

**Tabla 6.1. Checklist de construcción de indicadores múltiples**

| Etapa | Nombre | Descripción |
|---|---|---|
| **1** | **Marco teórico** | Proporciona la base para la selección y combinación de indicadores y variables en factores en un indicador compuesto (*subyacente*) significativo bajo el principio de ajuste (se involucra el conocimiento teórico-empírico en esta etapa). Es necesario realizarlo, para:<br>• Obtener una comprensión y definición clara del fenómeno multidimensional a medir.<br>• Para estructurar los distintos subgrupos del fenómeno (si fuera necesario).<br>• Para compilar una lista de criterios de selección para las variables subyacentes, por ejemplo: entrada, salida, proceso. |
| **2** | **Selección de datos** | Debe ser basado en la solidez analítica, mensurabilidad, ser representativa de la población a partir de la muestra, con relevancia de los indicadores relacionados entre sí y al fenómeno que deberá ser medido. Se consideran técnicas de sustitución y otras, cuando hay escasez de datos. Es necesario realizarlo, para:<br>• Comprobar la calidad de los indicadores disponibles.<br>• Verificar fortalezas y debilidades de cada uno de los indicadores.<br>• Crear una tabla resumen de características de datos, por ejemplo, disponibilidad (en el país, tiempo, etcétera), fuente, tipo (*hard, soft*, de entrada, de proceso, de salida, etcétera). |
| **3** | **Imputación de datos perdidos** | Su importancia radica en el hecho de que el conjunto de datos ausentes, debe justificarse. Es necesario, para:<br>• Estimar valores perdidos.<br>• Proporcionar una medida de la confiabilidad de cada valor imputado, así que evalúa el impacto de resultados de la imputación en el indicador compuesto (factor subyacente).<br>• Discutir la presencia de valores atípicos, así como técnicas de corrección en el conjunto de datos. |

| | | |
|---|---|---|
| 4 | **Análisis multivariado** | Debe usarse para estudiar la estructura general del conjunto de datos, evaluar su idoneidad, y seleccionar la técnica multivariante a utilizar (por ejemplo, ponderación, agregación, etcétera). Es necesario, realizarlo, para:<br>• Verificar la estructura subyacente de los datos a través de los indicadores utilizando técnicas como el análisis factorial exploratorio, el análisis de clúster, etcétera.<br>• Identificar grupos de indicadores que sean estadísticamente similares entre países e incluso, hacer una interpretación de los resultados.<br>• Comparar estadísticamente una determinada estructura del conjunto de datos con el marco teórico de la investigación y discutir las posibles diferencias. |
| 5 | **Normalización** | Debe llevarse a cabo para hacer que las variables sean comparables. Es necesario realizarlo, para:<br>• Seleccionar el procedimiento de normalización adecuada, que considere al marco teórico de la investigación como a las propiedades de los datos.<br>• Discutir la presencia de datos atípicos, ya que pueden convertirse en puntos de referencia no intencionados.<br>• Hacer ajustes de escala, si necesario.<br>• Transformar indicadores altamente sesgado, de ser necesario. |
| 6 | **Ponderación y agregación** | Debe hacerse de acuerdo con las líneas del marco teórico de la investigación. Es necesario realizarlo, para:<br>• Seleccionar la ponderación adecuada y el procedimiento de agregación que considere tanto lo teórico marco de la investigación como las propiedades de datos.<br>• Discutir sin los problemas de correlación entre los indicadores deben ser tomados en cuenta o no.<br>• Discutir si la compensación entre los indicadores se debe permitir. |
| 7 | **Análisis de incertidumbre y sensibilidad** | Debe realizarse para evaluar la robustez del índice o factor subyacente en términos de, por ejemplo, el mecanismo para incluir o excluir un indicador, el esquema de normalización, la imputación de datos ausentes, la elección de los pesos, el método de agregación. Es necesario realizarlo, para:<br>• Considerar una aproximación de modelado múltiple al construir el índice o factor subyacente, y si está disponible, escenarios conceptuales alternativos por la selección de los indicadores subyacentes.<br>• Identificar todas las fuentes posibles de incertidumbre en el desarrollo del índice o factor subyacente y determinar los límites de incertidumbre.<br>• Realizar análisis de sensibilidad de la inferencia (supuestos) y determinar qué fuentes de la incertidumbre son más influyentes en el puntajes y / o rangos. |

| | | |
|---|---|---|
| **8** | **Regreso a los datos** | Se debe asegurar que los datos sean los insumos que requiere el marco teórico de la investigación. La objetividad es fundamental para continuar los análisis que se requieran de los datos. Es necesario realizarlo, para:<br>• Perfilar el desempeño a nivel indicador<br>• Verificar la correlación y su causalidad (de ser posible).<br>• Identificar cómo, qué grupos y sub-grupos de indicadores están influyendo en el resultado del índice o factor subyacente |
| **9** | **Enlaces a otros indicadores** | Debe hacerse para correlacionar el índice o factor subyacente (o sus dimensiones) con indicadores existentes (simples o compuestos) así como para identificar enlaces a través de regresiones. Es necesario realizarlo, para:<br>• Correlacionar el índice o factor subyacente con otras medidas relevantes, tomando en consideración los resultados del análisis de sensibilidad.<br>• Desarrollar narrativas basadas en datos y resultados. |
| **10** | **Visualización de los resultados** | La visualización puede influir (o ayudar a mejorar) la interpretabilidad. Es necesario realizarlo, para:<br>• Identificar un conjunto coherente de herramientas de presentación para el público objetivo.<br>• Seleccionar la técnica de visualización que comunique más la información.<br>• Presentar el resultado del índice o factor subyacente lo más claro posible. |

Fuente: OECD (2008) con adaptación propia.

## Pros y Contras de usar indicadores

Los indicadores compuestos (que conforman a los factores subyacentes), que comparan el desempeño de innovación de un clúster de empresas o de un país son cada vez más reconocidos como herramienta útil de análisis de políticas y comunicación pública **(OECD, 2008)**. El número de indicadores compuestos (y por lo tanto, factores subyacentes) existentes alrededor del el mundo está creciendo año tras año **(Bandura, 2006**, cita más de **160** indicadores compuestos de todo tipo; **Dziallas y Blind (2019)** citan **82** tan solo para la innovación de procesos).

Tales indicadores compuestos (que conforman a los factores subyacentes) proporcionan comparaciones simples de agrupaciones de empresas

o países, que pueden ser utilizados para ilustrar problemas complejos y en ocasiones, difíciles de comprender en campos muy diversos de la innovación como los son: la innovación de productos/ servicios, la innovación por mercadotecnia, la innovación por tecnología, la innovación por el modelo de negocios, la innovación organizacional, la innovación social, etcétera. A menudo, parece más fácil para los investigadores, interpretar los indicadores compuestos (factores subyacentes) más que identificar tendencias comunes en muchos de los indicadores por separado, ya que también han demostrado ser útiles en la evaluación comparativa desempeño de un país **(Saltelli, 2007)** o de diversos tipos de organización. Sin embargo, los indicadores compuestos (factores subyacentes) pueden enviar políticas engañosas mensajes si están mal construidos o mal interpretados. Sus resultados de panorama general, pueden invitar a los investigadores para emitir conclusiones analíticas simplistas. De hecho, los indicadores compuestos (factores subyacentes) deben ser vistos como un medio para iniciar la discusión y estimular el interés público, tomando en cuenta el contexto. En términos generales, un indicador es una medida cuantitativa o cualitativa derivada de una serie de hechos observados que pueden revelar posiciones relativas (por ejemplo, de una organización como una pyme o un país) en un área y contexto determinados **(OECD, 2008)**. Cuando se evalúa en intervalos regulares, un indicador puede señalar la dirección del cambio a través de diferentes unidades y a través del tiempo. También pueden ser útiles para establecer prioridades políticas y para realizar evaluaciones comparativas o monitorear una intervención.

Un indicador compuesto (*factor subyacente*) se forma, cuando los indicadores individuales se compilan en un solo índice sobre la base de un modelo subyacente **(Mejía-Trejo, 2019b)**. El indicador compuesto (*factor subyacente*) idealmente debería medir la multidimensional de conceptos que no pueden ser capturados por un solo indicador, por ejemplo, la competitividad, la industrialización, la sostenibilidad, la integración en el mercado único, la sociedad basada en el conocimiento, etcétera. Los indicadores compuestos (*factor subyacente*) son muy parecidos a los modelos matemáticos. Como tal, su construcción se debe más a la capacidad de diseño

del modelador que a las reglas científicas universalmente aceptadas para codificar **(OECD, 2008)**. Con respecto a los modelos, la justificación de un indicador compuesto (*factor subyacente*) radica en su idoneidad para el propósito previsto y en la aceptación de pares **(Rosen, 1991)**. Sobre la disputa sobre si los indicadores compuestos (factor subyacente), son buenos o malos *per se*, se toma en cuenta por ejemplo, lo comentado por:

1. **Sharpe (2004):**

> *"...los agregadores creen que hay dos razones principales por las que es valioso combinar indicadores para que, de alguna manera, se produzca un resultado final. Creen que una estadística resumida pueda capturar la realidad y darle sentido, y que enfatizar el resultado final es extremadamente útil para obtene el interés de los medios y, por lo tanto, la atención de los responsables políticos. La segunda escuela, los no agregadores, cree que uno debería detenerse una vez que se haya creado un conjunto apropiado de indicadores y no ir más allá paso de producir un índice compuesto. Su objeción clave a la agregación es lo que ven como el naturaleza arbitraria del proceso de ponderación mediante el cual se combinan las variables..."*

2. **Saisana et al. (2005):**

> *"...es difícil imaginar que el debate sobre el uso de indicadores compuestos se resolverá alguna vez [...] los estadísticos oficiales pueden tender a resentir los indicadores compuestos, ya que existe mucho trabajo en la recolección de los datos y revisión que se desperdicia u oculta detrás de un número único de significado dudoso. Por otro lado, la tentación de los interesados y profesionales para resumir procesos complejos y a veces evasivos, como por ejemplo, la sostenibilidad, la política de mercado único, etc. en una sola figura como desempeño de referencia del país para el consumo de polí-ticas parece igualmente irresistible..."*

Este documento no tiene como objetivo resolver el debate, sino solo contribuir a una mejor comprensión de la *complejidad* de los indicadores compuestos (factores subyacentes) así como de referenciar al lector de sus las técnicas utilizadas (para mayor referencia, consultar **Mejía-Trejo (2019a y 2019b)** actualmente para construirlos. Ver **Tabla 6.2.**

**Tabla 6.2. Pros y Contras de diseñar y aplicar indicadores**

| Pros | Contras |
|---|---|
| • Puede resumir realidades complejas y Multidimensionales, para apoyo de toma de decisiones.<br>• Son más fáciles de interpretar que una batería de muchos indicadores separados.<br>• Sirve de base de evaluación del progreso de organizaciones como empresas y o países a lo largo del tiempo.<br>• Reduce el tamaño visible de un conjunto de indicadores, sin soltar la información subyacente de base.<br>• Hace posible incluir más información dentro del límite de tamaño existente.<br>• Ubica los problemas de desempeño de una organización como una empresa o país, de una forma clara.<br>• Facilitar la comunicación con el público en general. (es decir, ciudadanos, medios de comunicación, investigadores, especialistas, etcétera) promoviendo el monitoreo.<br>• Ayuda a construir y apuntalar narrativas para todo tipo de audiencia especializada y no especializada.<br>• Permite a los usuarios, comparar dimensiones complejas de manera efectiva. | • Puede enviar mensajes engañosos si es mal construido o mal interpretado.<br>• Puede invitar a conclusiones simplistas.<br>• Puede ser manipulado en sentido negativo, por ejemplo, para apoyar un cierto interés, si el proceso de construcción no es transparente y/o carece de un sólido marco teórico y empírico.<br>• La selección de indicadores y pesos podría ser objeto de conflicto de intereses.<br>• Puede ocultar fallas graves en algunas dimensiones de variables y aumentar la dificultad de identificación de medidas correctivas adecuadas, si el proceso de construcción no es transparente.<br>• Puede conducir a políticas inapropiadas si las dimensiones de rendimiento, que son difíciles de medir se ignoran. |

Fuente: OECD (2008) con adaptación propia

## SEM y fsQCA

Quizás, la implementación más sofisticada del enfoque de modelos teóricos basados en indicadores, es a través de la técnica analítica conocida como modelado por ecuación estructural (**SEM**. *Structural Equation Modeling*) **(Bollen, 1989; Mejía-Trejo, 2019b)**. **SEM** extiende el uso de múltiples indicadores de un concepto único (por ejemplo, el caso mostrado por **Mejía-Trejo et al., 2016)** a múltiples conceptos e interrelaciones. En esencia, el uso de **SEM** permite la construcción de indicadores compuestos o factores subyacentes con múltiples indicadores teniendo lugar, dentro del contexto del análisis de las interrelaciones entre conceptos. Por lo tanto, la construcción y ajuste del modelo del índice o de factores subyacentes, optimiza las relaciones hipotéticas y confirmar la validez y confiabilidad del instrumento diseñado.

Usando **SEM**, los investigadores pueden evaluar la coherencia de dichos indicadores compuestos o factores subyacentes construídos dentro del contexto del modelo en el que están implicados y simultáneamente, pueden evaluar la coherencia del modelo en su conjunto. El requisito clave, que debe cumplir un indicador es que debe ordenar los casos de acuerdo a sus cargas factoriales, de manera que se muestre el indicador compuesto o factor subyacente. Todas las técnicas aplicadas que plantean múltiples indicadores comparten una varianza observada, que a su vez es específica de la muestra en su definición y construcción.

El requisito clave que debe cumplir el planteamiento de indicadores, es que debe ordenar sus cargas factoriales, de manera que muestren el indicador compuesto o *factor subyacente* clasificándolos en altas o bajas cargas factoriales, para una muestra específicamente seleccionada.

Este tipo de clasificación entre cargas factoriales altas y bajas pueden ser consideradas medidas que ofrecen una forma de calibración rudimentaria y pasiva, incluso a partir de considerar que los datos solo cumplan con la condición de la linealidad, requerida como mínima en la aplicación de las técnicas multivariantes y particularmente en el **SEM,** es decir, sin considerar que el conjunto de datos y la muestra no cumplan con la normalidad o la *homocedasticidad* requeridas **(Hair et al., 1999; Mejía-Trejo, 2019a**

**y 2019b)** requerida por lo que varían de una muestra representativa que es la base del análisis multivariante. Por el contrario, en la práctica de las ciencias duras como en el caso de la química, los instrumentos de medición, antes deben ser calibrados con elementos externos (por ejemplo, la calibración de un termómetro en grados Celsius sobre los puntos de congelación del agua destilada a cero y su punto de ebullición a 100 grados Celsius), a fin de que sus resultados muestren confiabilidad y validez. A primera vista, estas prácticas convencionales con respecto al uso de indicadores en las ciencias sociales con orientación a la administración de la innovación, parecen ser completamente sencillos y no controvertidos. Por ejemplo, se aprecia completamente razonable, que el reporte *Global Innovation Index* (GIN, 2019) (ver **Figura 6.1**).

El cual propone que la innovación de un país (basados en indicadores compuestos que muestran al factor subyacente: innovación), se compone de dos variables y siete indicadores con los cuales, los países deben clasificarse entre sí y que alguna medida de tendencia central, basada en la muestra o población en cuestión, debe usarse para definir puntajes altos vs. bajos. De nuevo, el requisito fundamental del modelo teórico basado en indicadores, es simplemente la variación, que a su vez solo requiere:

- Una muestra (o población) que en donde descubrir la variedad de puntajes, y
- Una medida de tendencia central basada en la muestra (o población).

Tenga en cuenta, sin embargo, que en esta vista toda variación se considera igualmente relevante. Por supuesto, los investigadores tienen la opción de transformar sus variables, por ejemplo, usando transformaciones logarítmicas en la preparación de los datos **(Mejía-Trejo, 2019a**, Capítulo 3) para reducir la inclinación y cambiar el peso de la variación. Sin embargo, tales ajustes son relativamente poco frecuentes y, en cualquier caso son, entendidos como como una forma de mejorar la solidez de un modelo. Es decir, la variación en todo el rango del indicador, se considera pertinente con respecto a lo que se revela sobre del factor subyacente.

**Figura 6.1. Modelo conceptual de la Global Innovation Index**

INDICADORES | VARIABLES | FACTORES

INSTITUTIONS
Political environment
Regulatory environment
Business environment

HUMAN CAPITAL AND RESEARCH
Education
Tertiarty education
Research & development (R&D)

INFRASTRUCTURE
ICTs
General infrastructure
Ecological sustainability

MARKET SOPHISTICATION
Credit
Investment
Trade, competition, & market scale

BUSINESS SOPHISTICATION
Knowledge workers
Innovation linkages
Knowledge absorption

KNOWLEDGE AND TECHNOLOGY OUTPUTS
Knowledge creation
Knowledge impact
Knowledge difusion

CREATIVE OUTPUTS
Intangible assets
Creative goods and services
Online creativity

Innovation Input Sub-Index

Innovation Output Sub-Index

FACTOR SUBYACENTE

GLOBAL INNOVATION INDEX

Fuente: GII (2019).

Por ejemplo, los países que se colocan en la parte superior del *Global Innovation Report Index* **(GII, 2019)** son marcados como países altamente innovadores, de acuerdo a la categoría de ingresos (*High, Upper-middle, Low-middle or Low incomes*) (ver **Tabla 4.1**). Sin embargo, la diferencia que los separa indica que uno es aún más altamente desarrollado que el otro. En el enfoque del modelo teórico basado en indicadores, esta diferencia generalmente se toma al pie de la letra, lo que significa que no se intentan observar los casos y preguntar si esta diferencia o cualquier otra diferencia, independientemente de la magnitud, es relevante o significativa con respecto al indicador compuesto o factor subyacente.

Observe también, que la idea de que la variación en cualquier extremo de una distribución, debería ser desalentada de alguna manera, es usualmente vista con gran sospecha por los investigadores cuantitativos, porque suprimir dicha variación tiende a atenuar a las correlaciones. Por el contrario, la interpretación de los puntajes en relación con los estándares externos acordados, es central para la calibración de los instrumentos de medición. Estas normas externas proporcionan un contexto para la interpretación de las puntuaciones.

## La investigación cualitativa: prácticas de mediciones comunes

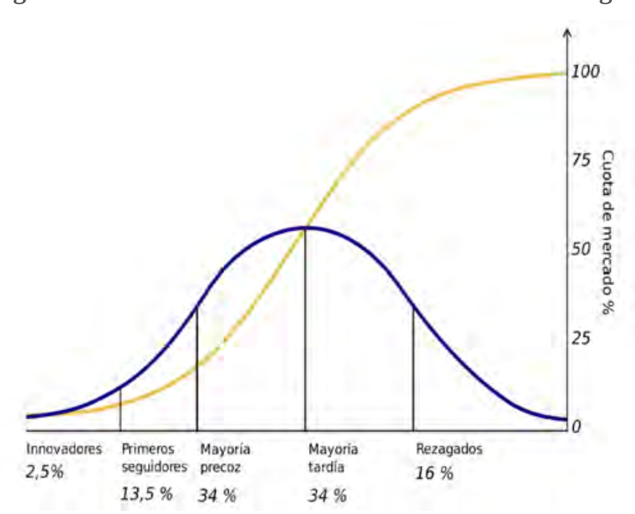
En la investigación cuantitativa convencional, las mediciones que se realizan con los indicadores son sobre conceptos derivados de los modelos sustentados en los marcos teóricos de una investigación. Por lo tanto, la investigación con enfoque cuantitativo de medición, está fuertemente centrado en la teoría. Por el contrario, mucho de la investigación con enfoque cualitativo de medición, es más bien basada en la experiencia de los sujetos, por lo tanto, centrado en evidencia empírica y por lo tanto, de naturaleza más iterativa, por lo que es posible llamarlo, modelo basado en casos. Es decir, hay una interacción entre la formación del concepto y la medición por un lado, y la estrategia de investigación por el otro **(Glaser y**

**Strauss, 1967)**. El investigador comienza con un amplio planteamiento de ideas y conceptos y utiliza casos empíricos para ayudar a refinarlos **(Mejía-Trejo, 2019b)**. Este proceso de refinamiento progresivo implica un movimiento iterativo de análisis y síntesis entre ideas y evidencia **(Katz, 1982; Ragin, 1994)**, donde los investigadores especifican y refinan la forma de realizar las mediciones así como los indicadores empíricos.

## Teoría de difusión de innovación de Rogers y Ley de Moore

Considere por ejemplo la teoría de difusión de la innovación de Rogers **(Rogers, 1962)**, la cual es una teoría sociológica que pretende explicar cómo, por qué y a qué velocidad se mueven las nuevas ideas (y tecnologías) a través de las diversas culturas. El principal elemento, la innovación que es definida por **Rogers (1962)** como una idea, práctica u objeto que es percibido por un individuo como nuevo. En la teoría de difusión se distingue entre los individuos que aceptan las innovaciones en los primeros instantes de su emisión y aquellos que lo hacen en etapas posteriores. En este sentido, se establecen cinco categorías de adoptantes en función del tiempo que requieren los individuos para adoptar una innovación. Las categorías de adoptantes son: los innovadores, los primeros adoptantes, la mayoría precoz, la mayoría rezagada y los tradicionales. Ver **Figura 6.2.**

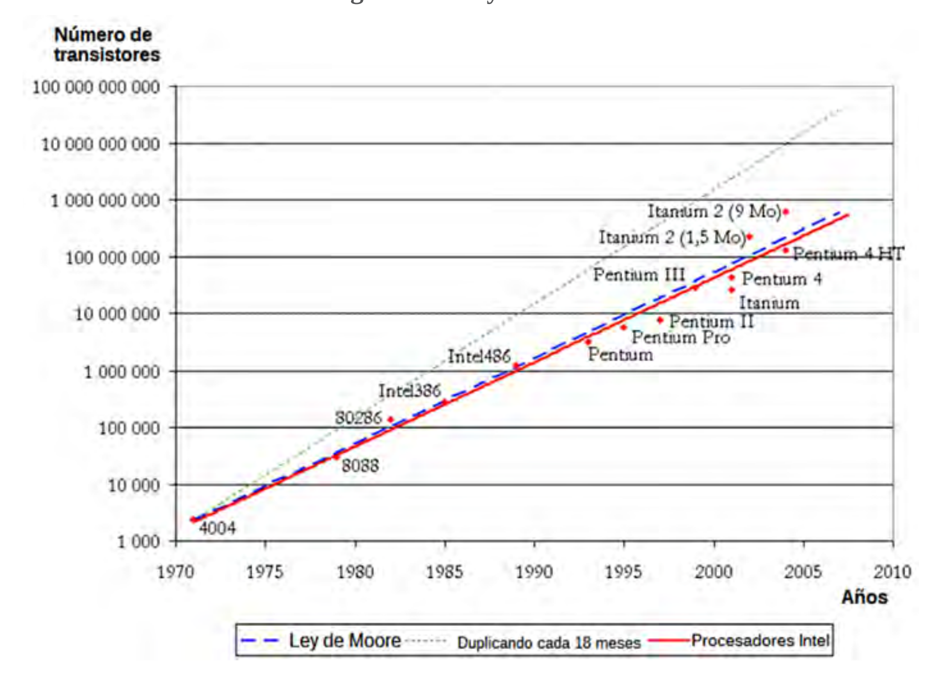
**Figura 6.2. Curvas de difusión de la innovación de Rogers**




Fuente: Rogers (1962) con adaptación propia.

Sin embargo, ¿qué es la formación de cada una de estas categorías? Por ejemplo, como se aprecia, las categorías se representan a partir de un porcentaje de cuota de mercado, pero también, podría ser fechado y desarrollar mediante alguna escala de razón una medición de edad de la innovación, que pudiera ser relevante para el concepto de formación de innovaciones. Por ejemplo, actualmente en la industria de los microprocesadores, se tiene en cuenta la Ley de Moore, emitida por Gordon Moore (ex CEO de Fairchild Semiconductor) y quien el 19 de abril de 1965, la revista *Electronics* publicó un documento, en el que él anticipaba que la *complejidad* de los circuitos integrados se duplicaría cada año con una reducción de costo conmensurable**.** Ver **Figura 6.3.**

**Figura 6.3. Ley de Moore**




Fuente: Cnet (2011),

Cabe destacar que la Ley de Moore no es una ley en estricto sentido científico, se le considera una ley empírica basada en la observación, pero ha sentado las bases de grandes saltos de progreso. Ya sea tasa de mercado (porcentajes) o tiempo (meses, años) deben ajustarse en la base del conocimiento sustantivo acumulado para poder interpretar temprano versus tarde de una manera que concuerde apropiadamente con el marco teoríco que lo sustente.

## La necesidad de calibrar

Ambos casos presentados **Rogers, (1962)** y **Moore (1965)** requerirían de calibrar sus mediciones. Tales calibraciones deberían ser rutinarias en el enfoque de investigación cualitativo, aunque rara vez se modelan o incluso

se declaran explícitamente **(Ragin, 2008)** de hecho, desde el enfoque de la investigación cuantitativa convencional, parece que los investigadores cualitativos sesgan sus mediciones para que se ajusten a sus ideas preconcebidas y simplemente interpretar tales indicadores. Una segunda característica esencial de la medición en la investigación cualitativa es que está más orientado a casos que la medición en la investigación cuantitativa.

En la investigación cualitativa orientada a casos, el enfoque conceptual está basado en las categorías específicas de los casos, por ejemplo, las empresas y/o países innovadores. En la investigación por modelo teórico basado en indicadores, por el contrario, la atención se centra en las dimensiones de variación de indicadores, variables o factores que se presentan en una determinado muestra o población de casos. La distinción es sutil, pero importante porque los casos pueden variar no solo a lo largo de los indicadores de qué tan bien satisfacen los requisitos para ser miembro de una categoría o conjunto.

Por ejemplo, los países que se presentan como innovadores en el reporte de la *Global Innovation Index* **(GII, 2019)**, varían en qué tan bien satisfacen los requisitos de pertenencia en el conjunto de países innovadores: algunos casos los satisfacen completamente, otros parcialmente y algunos no del todo. A fin de evaluar qué tan bien los casos satisfacen los requisitos de pertenencia, es necesario invocar otros índices externos para convalidar las categorías del factor subyacente. Observe por cierto, la **Tabla 6.2,** basada en la **OECD (2008)** por ejemplo, los indicadores de referencia de calibración, no son mencionados.

Así, en el enfoque de investigación cualitativa orientada a casos, el foco principal está en el conjuntos de casos, es decir, cada uno de los miembros de los cuales se pueden identificar y estudiar individualmente (por ejemplo, los países innovadores con ingresos medios (*Upper Middle-Income*) en la línea de expectativas del nivel de desarrollo (*In line with expectations with level development*) de la **Tabla 4.1.**

**Tabla 4.1. Caso de México de acuerdo al Global Innovation Index (2019)**

**Innovation performance at different income levels, 2019**

| | High Income | Upper-middle Income | Lower-middle Income | Low Income |
|---|---|---|---|---|
| Above expectations for level of development | Denmark | Armenia | Georgia | Burundi |
| | Finland | China | India | Malawi |
| | Netherlands | Costa Rica | Kenya | Mozambique |
| | Singapore | Montenegro | Mongolia | Rwanda |
| | Sweden | North Macedonia | Philippines | Senegal |
| | Switzerland | South Africa | Republic of Moldova | United Republic of Tanzania |
| | United Kingdom | Thailand | Ukraine | Tajikistan |
| | United States of America | Malaysia | Viet Nam | Uganda |
| | Germany | Bulgaria | Tunisia | Nepal |
| | Israel | Romania | Morocco | Ethiopia |
| | Republic of Korea | Mexico | Indonesia | Mali |
| | Ireland | Serbia | Sri Lanka | Burkina Faso |
| | Hong Kong, China | Iran (Islamic Republic of) | Kyrgyzstan | Madagascar |
| | Japan | Brazil | Egypt | Zimbabwe |
| | France | Colombia | Cambodia | Niger |
| In line with expectations for level of development | Canada | Peru | Côte d'Ivoire | Benin |
| | Luxembourg | Belarus | Honduras | Guinea |
| | Norway | Bosnia and Herzegovina | Cameroon | Togo |
| | Iceland | Jamaica | Pakistan | Yemen |
| | Austria | Albania | Ghana | |

Fuente: GII (2019).

Por el contrario, en el modelo teórico basado en indicadores, los casos generalmente se entienden simplemente como muestras de donde tomar mediciones, que a su vez proporcionan la materia prima necesaria para estudiar las relaciones entre indicadores, variables o factores. Es posible por tanto afirmar, qué enfoque de investigación cualitativa orientada a casos es más compatible con la idea de que las mediciones se deben calibrar, ya que el enfoque está orientado al grado a qué casos satisfacen los criterios de pertenencia, que a su vez suelen estar referenciados a indicadores externos, que no son producto del análisis multivariante. Estos criterios de pertenencia, deben reflejar los estándares acordados, de lo contrario, se pondrá en duda la constitución de una categoría o conjunto **(Ragin, 2008)**. En el enfoque del modelo teórico basado en indicadores, los miembros de una población simplemente varían en el grado en que expresan un rasgo o fenómeno dado, y generalmente no hay una motivación especial para especificar los criterios para ser el sujeto en estudio, un miembro de un

conjunto o para identificar casos específicos como instancias. En resumen, en la investigación cualitativa se interpreta.

**Cicourel (1964)** argumenta que es necesario considerar los tres medios a través de los cuales los científicos cualitativos desarrollan categorías y las vinculan a propiedades observables de objetos y eventos, que son: lenguaje, significado cultural y las propiedades de los sistemas de medición. Basados en este argumento, el problema de establecer equivalencia de categorías (como países innovadores o no innovadores) no deben ser una medición única o descriptiva ya que debe considerar los problemas del lenguaje y el significado cultural. De hecho afirma **(Cicourel, 1964, p. 33)**:

> *"...considerar las variables como cuantitativas sólo porque los datos disponibles se expresan en forma numérica o porque se considera más –científico– no proporciona una solución a los problemas de medición .... no sustituye el examen y reexaminando la estructura de nuestras teorías para que nuestras observaciones, descripciones y medidas de las propiedades de los objetos sociales y los eventos tengan una correspondencia literal con lo que creemos ser la estructura de la realidad social..."*

En términos simples, Cicourel (1964) argumenta que las medidas y sus propiedades deben evaluarse en el contexto de ambos conocimientos: el teóricos y el empírico. El hecho de que los científicos sociales posean un indicador de escala de razón de un concepto teórico, no significa que este aspecto de la realidad social tenga las propiedades matemáticas de este tipo de escala.

Por lo tanto, en la investigación cualitativa, la idea de que los científicos sociales deban utilizar estándares externos para evaluar e interpretar sus mediciones tiene mucho más mérito que lo realizado en la investigación cuantitativa convencional. Una diferencia importante con la investigación cuantitativa, sin embargo, es que la medición en la investigación cualitativa generalmente carece de precisión, y la forma de medir la sensibilidad al contexto, orientada a casos, rasgo típico de la investigación cualitativa, a menudo parece casual y poco científico **(Ragin, 2008)**.

## El conjunto de datos difusos (*fuzzy-sets*) como puente entre ambos enfoques

Con los conjuntos de *datos difusos*, es posible tener lo mejor de ambos enfoques (*cuantitativo y cualitativo*), es decir, el uso de conocimiento teórico con precisión, que es apreciado por los investigadores cuantitativos y el uso de conocimiento empírico sustantivo para calibración de las mediciones, por los investigadores cualitativos. La precisión proviene, en forma de cuantitativa, de las evaluaciones del grado de pertenencia al conjunto de datos, que pueden ir desde una puntuación de **0.0** (*exclusión completa de un conjunto*) a **1.0** (*inclusión completa*).

El conocimiento empírico sustantivo, proporciona los criterios externos que lo hacen posible, para calibrar las mediciones. Este conocimiento indica lo que constituye el ser miembro de *inclusión completa*, *la no pertenencia completa y el punto en el que los casos intermedios, los que están más dentro que fuera del conjunto* **(Ragin, 2000; Smithson y Verkuilen, 2006)**.

Los criterios externos, que se utilizan para calibrar las mediciones y traducirlos en puntuaciones de pertenencia al conjunto de datos, pueden reflejar las normas basadas sobre el conocimiento social para la innovación de un país. Por ejemplo, el Plan Nacional de Ciencia, Tecnología e Innovación **(PECITI)** de México, el cual se genera para un sexenio con una visión de largo plazo a veinticinco años, con revisión de cada tres años, en coincidencia con el inicio de cada nueva Legislatura del Congreso de la Unión **(PECITI, 2014-2018)**, o del propio conocimiento acumulado del investigador del que derive el estudio de casos concretos, por ejemplo, a la entrada de nuevas innovaciones basadas por diferentes tipos de obsolescencia, tales como: programada, indirecta, funcional por defecto, incompatibilidad, psicológica, estética, por notificación, caducidad, ecológica, etcétera. Estos criterios externos, deben estar descritos de manera clara y explícita, así también, deben ser aplicados de forma sistemática y transparente. Este requisito separa el uso del conjunto de *datos difusos*, del trabajo convencional cualitativo, donde los estándares que se aplican, por lo general, permanecen de forma implícita. De hecho, el conjunto de *datos*

*difusos*, trabaja simultáneamente con ambos enfoques: cualitativa y cuantitativa, considerado que son un puente ideal entre ellos, por:

1. La designación de pertenencia completa y no completa (estados **1.0** y **0.0**), los cuales son estados cualitativos y entre estos dos estados se encontrarán variando cualitativamente, los grados de pertenencia que van desde más fuera (más cerca de **0.0**) y más dentro (más cerca de **1.0**).
2. Aplicar ambos enfoques de investigación: modelo teórico basado en indicadores (*cuantitativo*) y el modelo basado en casos (*cualitativo*).
3. Están orientados a modelo basado en casos cuando el tipo o categoría de los casos es determinante; están orientados a modelo teórico basado en indicadores cuando se permite niveles o grados de pertenencia y por lo tanto a una graduación entre casos. Esta cualidad, provee precisión a la medición, lo cual es altamente apreciado en la investigación cuantitativa.

## El conjunto de datos difusos (*fuzzy-sets*) y su relación con las variables

Una diferencia clave entre un conjunto de *datos difusos* y una variable convencional, es la forma en que se conceptualizan y etiquetan. Por ejemplo, mientras que es posible construir una variable genéricas tales como nivel de educación de los CEOs, es asimismo, imposible transformar esta variable directamente a un conjunto de *datos difusos*, sin primero designar y definir, el objetivo de un conjunto de casos. Así, el investigador podría estar interesado en el conjunto de empresas innovadoras con al menos de CEOs de nivel de educación secundaria o tal vez el conjunto de individuos que tengan educación universitaria o de posgrado. Este ejemplo, pone de manifiesto que la designación de diferentes esquemas de calibración. Por ejemplo, para el caso de un CEO que tiene un año de educación universitaria, tendría plena adhesión o pertenencia equivalente a **(1.0),** en el conjunto de CEOs que tengan al menos educación secundaria, pero este mismo CEO tiene claramente la no pertenencia en el conjunto de CEOs

que son graduados universitarios. Tome en cuenta, que en el contexto de las empresas innovadoras, la variable nivel de educación de los CEOs, es genérica y que deberá calibrarse, si se requiere incorporar al conjunto de *datos difusos*, por ejemplo, en países innovadores con ingresos medios (*Upper Middle-income*) en la línea de expectativas del nivel de desarrollo (*In line with expectations with level development*) **(GII, 2019)**. Tenga en cuenta que el requisito que el investigador designe un objetivo, no solo estructura la calibración del conjunto de *datos difusos*, sino que también proporciona una conexión directa entre el discurso del marco teórico de la investigación y el análisis empírico. Después de todo, es más común para el discurso del marco teórico que se organiza en torno a los diferentes tipos o categorías de casos (por ejemplo, países innovadores) de lo que son para la prescripción de las variables genéricas (por ejemplo, nivel de educación de los CEOs). Estos ejemplos, aclaran una característica crucial de los conjuntos de *datos difusos* central a su calibración y es el el hecho, de que a fin lograrla, es necesario que los investigadores distinguan entre una variación *relevante e irrelevante*. Por ejemplo, en nuestro caso, la diferencia entre CEO que ha completado un año de la universidad y otro que ha completado dos años de universidad es *irrelevante* para el conjunto de los individuos con al menos una educación secundaria, tanto que ambos individuos pertenecen por completo al grupo (pertenencia = **1.0**). La diferencia de un año, simplemente no es relevante al conjunto objetivo como conceptualizado y etiquetado. Cuando se calibra un conjunto de *datos difusos*, la variación de los conceptos teóricos, se etiquetan que son *irrelevantes* para el conjunto truncado por las puntuaciones de pertenencia resultantes y que reflejan fielmente la composición de etiqueta del conjunto objetivo. Este requisito también se establece una estrecha relación entre el marco teórico y el análisis empírico. Una gran ventaja de utilizar cuidadosamente la calibración de conjuntos de datos difusos es, que permiten el uso de los principios del marco teórico en la investigación social orientada a la innovación. Estos principios incluyen:

1. El subconjunto de datos (que son centrales, para el análisis de *necesidad y suficiencia*),
2. La intersección de conjuntos (que es central para el estudio de configuración de casos),
3. La unión de conjuntos (que es fundamental para el examen de trayectorias alternativas hacia un mismo resultado),
4. Las *tablas de verdad* (que se utilizan para desentrañar la causa de la *complejidad*), y así sucesivamente.

Estas operaciones del marco teórico de la investigación están fuera de los límites a los investigadores que utilizan mediciones no calibrados, tales como las variables de escala tanto de intervalo como de razón.

Este capítulo explora la calibración de los conjuntos de *datos difusos* con mayor detalle, con una énfasis a lo práctico. Se centra en la calibración de las *variables de escala de intervalo y razón*, como conjuntos de *datos difusos* y se describirán en dos métodos generales:

- El *método directo*, que se basa en los puntos de referencia de investigación especificados como **1**= *completamente dentro* / **0**= *completamente fuera* / **0.5**= *ni completamente dentro ni completamente fuera*.
- El *método indirecto*, basado en la clasificación que los investigadores propongan sobre los casos, en seis categorías y el uso de un procedimiento de estimación de regresión para traducir las puntuaciones totales (*brutas*) en puntuaciones de pertenencia difuso.

## Técnicas de calibración de datos difusos (*fuzzy-sets*)

Se esbozan dos técnicas convencionales para la calibración de variables de escala de intervalo como conjunto de *datos difusos*, utilizando estándares externos para la calibración. Como se ha mencionado, las variables convencionales, ya sea sin calibrar o calibradas de forma implícita a partir del marco teórico, utilizando inducción derivada (uso de la media, desviación standard y varianza de la muestra poblacional). Por el contrario, los

conjuntos de *datos difusos*, se calibran mediante el uso de criterios externos, que a su vez deberá derivarse y ajustarse a la conceptualización, definición y etiquetado del conjunto en cuestión del investigador. Las normas externas para la calibración, se pueden implementar de dos maneras diferentes:

1. El primero es el uso del *método directo*, el investigador especifica los valores de escala de intervalo, que corresponden a los tres puntos mínimos de referencia del enfoque cualitativo, que se encuentran en un conjunto de *datos difusos*: **1**= completamente dentro / **0**= completamente fuera / **0.5**= ni completamente dentro ni completamente fuera (punto de traslape, superposición o *crossover*). Estos tres puntos de referencia, utilizan los valores originales de intervalos de escala y los transforman a las puntuaciones de pertenencia difusa.
2. El segundo, es el uso del *método indirecto*, donde el estándar externo que se utiliza es la evaluación cualitativa que realiza el investigador del grado en que los casos con puntuaciones otorgadas en una escala de intervalo, los hace miembros del conjunto objetivo. Los investigadores asignan cada caso en una de seis categorías y posteriormente aplica una simple técnica de estimación para re-escalar la medición original para que se ajuste a estos evaluaciones cualitativas. El producto final, de ambos métodos es la graduación de la calibración del grado de pertenencia de los casos en los conjuntos de *datos difusos*, con puntuaciones que van de **0.0** a **1.0**. Los ejemplos proporcionados en este capítulo, ilustran la capacidad de respuesta de estos métodos de calibración a la conceptualización del investigador al conjunto de datos objetivo.

## De variables de escala de intervalo a conjunto de datos difusos (*fuzzy-sets*)

Idealmente, la calibración del grado de pertenencia a un conjunto debe ser basado enteramente en el conocimiento tanto empírico sustantivo como de marco teórico que sustenta al investigador. Es decir, la base de conocimiento colectivo de los científicos sociales, debe ser la base de la

especificación para la calibración de mediciones precisas. Por ejemplo, los científicos sociales deben ser capaces de especificar diversos indicadores que describen la administración del conocimiento que propician la innovación como liderazgo, captura y adquisición de conocimiento, entrenamiento y mentoría, políticas y estrategias, comunicaciones, incentivos **(OECD, 2004)** y determinar los niveles de pertenencia de si existen organizaciones, empresas o países con prácticas de innovación. Sin embargo, las ciencias sociales en general, se encuentran todavía en etapas incipientes, y esta base de conocimientos aún no existe. Por otra parte, el predominio de investigación orientado a modelos téoricos basados en indicadores con sus referentes en la práctica del análisis multivariante como clave para evaluar las relaciones entre los casos socava el interés académico en determinar los umbrales puntos de referencia. Mientras que el problema de especificar umbrales y puntos referencia no ha ganado la atención que merece, se considera que no es una tarea de enormes proporciones. El requisito principal para lograr una calibración útil del instrumento de medición, se basa simplemente en atender cuáles son las cuestiones a resolver del objeto de estudio (por ejemplo, el establecimiento de los umbrales que constituyan la pertenencia: plena, no plena o zona de traslape, de cierto tipo el liderazgo como el transaccional en los CEOs de una empresa innovadora) (Ver **Mejía-Trejo et al., 2013**).

A pesar de las imperfecciones de la base de conocimiento existente, todavía es posible demostrar técnicas de calibración; todo lo que falta son estándares acordados, precisos para las mediciones calibradas. A manera demostrativa académica y en la medida de lo posible, las calibraciones presentadas en este capítulo se basarán en literatura teórica y empírica sustantiva existente. Aún así, se recuerda que el enfoque está más bien en las técnicas de calibración, y no en los puntos de referencia empíricos específicos utilizados para estructurar la calibración. Las técnicas presentadas suponen:

1. Que los investigadores ya tienen su disposición de indicadores basados en escala de intervalos convencionales de sus conceptos (*escala de Likert*).
2. Que el factor subyacente ha sido estructurado y etiquetado en términos de un marco teórico pre-existente. Tenga en cuenta que este requisito de etiquetado en las variables, mueve al modelo orientado a casos de investigación.
3. El *método directo*, utiliza estimaciones del registro de las probabilidades de plena pertenencia en un conjunto de datos, como un paso intermedio, por lo que su utilidad se hará evidente a medida que la avance la demostración de sus datos resultantes.

Por ahora, considere la **Tabla 6.3.** que muestra las diferente métricas que se utilizan en la demostración del *método directo*, donde:

**Tabla 6.3.Categorías cualitativas de 9 niveles**

| Nivel | Columnas | | | |
|---|---|---|---|---|
| | **1** | **2** | **3** | **4** |
| | **Etiqueta** | **Grado de pertenencia** | **Posibilidad de pertenencia (*odds of membership*)** | **Posibilidad de pertenencia por logaritmo natural** |
| | | **Y** | **P = (Y ) / (1-Y)** | **Ln (P)** |
| **1** | Plena pertenencia (*Full membership*) | 0.993 | 141.86 | **5.0** |
| **2** | Umbral de plena pertenencia (*Threshold of full membership*) | 0.953 | 20.28 | **3.0** |
| **3** | Principalmente dentro (*Mostly in*) | 0.881 | 7.40 | **2.0** |
| **4** | Más dentro que fuera (*More in than out*) | 0.622 | 1.65 | **0.5** |
| **5** | Punto de traslape (*Crossover point*) | 0.500 | 1.00 | **0.0** |
| **6** | Más fuera que dentro (*More out than in*) | 0.378 | 0.61 | **- 0.5** |
| **7** | Principalmente fuera (*Mostly out*) | 0.119 | 0.14 | **- 2.0** |
| **8** | Umbral de no plena pertenencia (*Threshold of full non-membership*) | 0.047 | 0.05 | **- 3.0** |
| **9** | No plena pertenencia (*Full non-membership*) | 0.007 | 0.01 | **- 5.0** |

Fuente: Ragin (2008) con adaptación propia.

1. La *columna 1*, muestra las varias etiquetas que pueden ser asignadas y que van desde pertenencia completa (*full membership*) hasta no pertenencia completa (*full non-membership*) para nueve niveles.
2. La *columna 2*, muestra el grado del conjunto de pertenencia vinculado a cada etiqueta. Para mayor facilidad, el grado de pertenencia es redondeado a tres decimales.
3. La *columna 3*, muestra las posibilidades de estar en plena pertenencia que resultan de la transformación de las puntuaciones de pertenencia del

conjunto (*columna 2*) en las posibilidades de caída pertenencia, utilizando la siguiente fórmula:

**P = (Y) / (1-Y)**

**4.** La *columna 4* muestra el logaritmo natural de las posibilidades reportadas en la *columna 3*. En efecto, las *columna 2 a 4* son diferentes representaciones del mismo valor numérico, utilizando diferentes métricas. Por ejemplo, la puntuación de pertenencia asignada a umbral de plena pertenencia (*Threshold of full membership*) es **0.953.**

La conversión a un valor posibilidades produce **20.28**. El cálculo del logaritmo natural de **20.28** produce un puntaje de **3.0**. Trabajar en esta métrica es útil, porque es completamente simétrica alrededor de **0.0** (o posibilidades de **50 / 50**) y no sufre de los efectos, ni piso ni techo de los valores. Así por ejemplo, si una técnica de calibración devuelve un valor en las posibilidades logarítmicas naturales de que sea, un valor muy grande ya sea positivo o negativo, su traducción al grado estancias de pertenencia estará dentro de los límites de **0.0** a **1.0**, que es un requisito básico de las puntuaciones de pertenencia difusa.

La tarea esencial de la calibración, usando el método directo, es transformar las variables de escala de intervalo en la métrica de posibilidad de pertenencia logarítmica natural en una manera que respete las etiquetas que se muestran en la *columna 1* de la **Tabla 6.3.** Los procedimientos para la calibración de las puntuaciones de pertenencia difusa, presentados en este capítulo, son matemáticamente incapaces de producir exactamente conjunto de pertenencias con puntuaciones de **1.0** o **0.0**. Estas dos puntuaciones de pertenencias, correspondería a infinito positivo y negativo, respectivamente, para el registro de las posibilidades. En cambio, las puntuaciones que son mayores que **0.95,** pueden ser interpretadas como de plena pertenencia de miembros en el conjunto objetivo, y las que son menores a **0.05** pueden ser interpretado como no plena pertenencia.

## Método directo de calibración

El punto de partida de cualquier calibración es la clara especificación del objetivo. El objetivo de esta demostración, es el hipotético caso de un clúster de 136 empresas de un clúster, altamente competitivas de alta tecnología.

### Caso 1

La **Tabla 6.4** muestra los datos de las **24** más relevantes de estos **136** empresas del cluster, con la siguiente pregunta:

**P1.** ¿Cómo calibrar los datos de la **Tabla 6.4.** basados en sus ingresos anuales por innovación, para que sean insumos de investigación, con un punto de cruce = **5000**; superior al punto de cruce= **20000**; inferior al punto de cruce= **2500**?

El *método directo* se realiza, como sigue:

1. Utiliza tres importantes referencias cualitativas para realizar la calibración: el umbral para la *plena pertenencia, el umbral de no plena pertenencia y el punto de cruce o traslape* .
2. El *punto de cruce o traslape* es el valor de una variable de la escala de intervalo, donde hay una gran ambigüedad en cuanto a si un caso está más dentro o más fuera de la meta establecida. Para los efectos de esta demostración, los *ingresos anuales por innovación* (mmusd) a utilzar será de un valor **5000** (mmusd) como *punto de cruce* y se muestran en la *columna 1*.
3. Un paso importante en el método directo de calibración, es calcular las desviaciones de puntaje bruto, que se muestra en la *columna 2* como desviaciones del *punto de cruce* designado por el investigador (**5000** este ejemplo). Estos valores se muestran en la *columna 2* de la **Tabla 6.4**. Las puntuaciones negativas, indican que un caso está más fuera que dentro del conjunto de datos objetivo, mientras que puntuaciones positiva de que el caso está más dentro que fuera.

**Tabla 6.4. Caso hipotético: calibrando el grado de pertenencia en las empresas competitivas de alta tecnología clúster. Método directo**

| Empresas competitivas de alta tecnología clúster | Columna | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| | Ingresos anuales por innovación (mmusd) | Desviaciones del punto de cruce (crossover) | Escalares Puntos Cruce=5000 Superior=20000 Inferior= 2500 | Posibilidad de pertenencia por logaritmo natural | Grado de pertenencia | |
| | I | D = ( I – 5000 ) | X 3 / 20000= 0.0002 3 / 2500= 0.0012 | P = (D) * (X) | Y = [exp (P) ] / [ 1+ exp (P)] | |
| | | | | | Y | Y Redondeo |
| 1 | 40110 | 35110 | 0.0002 | 7.02 | 0.99911 | 1.00 |
| 2 | 34400 | 29400 | 0.0002 | 5.88 | 0.99721 | 1.00 |
| 3 | 25200 | 20200 | 0.0002 | 4.04 | 0.98271 | 0.98 |
| 4 | 24920 | 19920 | 0.0002 | 3.98 | 0.98166 | 0.98 |
| 5 | 20060 | 15060 | 0.0002 | 3.01 | 0.95302 | 0.95 |
| 6 | 17090 | 12090 | 0.0002 | 2.42 | 0.91834 | 0.92 |
| 7 | 15320 | 10320 | 0.0002 | 2.06 | 0.88695 | 0.89 |
| 8 | 13680 | 8680 | 0.0002 | 1.74 | 0.85069 | 0.85 |
| 9 | 11720 | 6720 | 0.0002 | 1.34 | 0.79249 | 0.79 |
| 10 | 11290 | 6290 | 0.0002 | 1.26 | 0.77903 | 0.78 |
| 11 | 10940 | 5940 | 0.0002 | 1.19 | 0.76674 | 0.77 |
| 12 | 9800 | 4800 | 0.0002 | 0.96 | 0.72312 | 0.72 |
| 13 | 7470 | 2470 | 0.0002 | 0.49 | 0.62011 | 0.62 |
| 14 | 4670 | - 330 | 0.0012 | -0.40 | 0.40131 | 0.40 |
| 15 | 4100 | - 900 | 0.0012 | - 1.08 | 0.25351 | 0.25 |
| 16 | 4070 | - 930 | 0.0012 | - 1.12 | 0.24601 | 0.25 |
| 17 | 3740 | - 1260 | 0.0012 | - 1.51 | 0.18094 | 0.18 |
| 18 | 3690 | - 1310 | 0.0012 | - 1.57 | 0.17222 | 0.17 |
| 19 | 3590 | - 1410 | 0.0012 | - 1.69 | 0.15578 | 0.16 |
| 20 | 2980 | - 2020 | 0.0012 | - 2.42 | 0.08166 | 0.08 |
| 21 | 1000 | - 4000 | 0.0012 | - 4.80 | 0.00816 | 0.01 |
| 22 | 650 | - 4350 | 0.0012 | - 5.22 | 0.00538 | 0.01 |
| 23 | 450 | - 4550 | 0.0012 | - 5.46 | 0.00424 | 0.00 |
| 24 | 110 | - 4890 | 0.0012 | - 5.87 | 0.00281 | 0.00 |

Notas: con propuestas de punto de cruce = 5000; superior al punto de cruce= 20000; inferior al punto de cruce= 2500; numerador 3, ver cálculo en los puntos 7,8 y 12 abajo enunciados.

Fuente: Elaboración propia.

4. Para el *umbral de la plena pertenencia (threshold for full membership*) en el conjunto objetivo, usted podría determinar utilizar *ingresos anuales por innovación* (mmusd) de **20000**, que tendría una puntuación de desviación de **15000** (comparar las *columnas 1 y 2* de la **Tabla 6.4**). Este valor, corresponde a establecer una puntuación de pertenencia **0.95** (ver **Tabla 6.3**) con una *posibilidad logarítmica natural* de **3.0**. Por lo tanto, los casos con los *ingresos anuales por innovación* (mmusd) de **20000** o mayores (es decir, la desviación de **15000** o mayores) se consideran de plena pertenencia, con puntuaciones **>= 0.95** y posibilidad logaritmica natural de pertenencia de **>= 3.0**
5. Por el contrario, en la dirección inversa, el umbral para la no plena pertenencia (*threshold for full non-membership*) en el conjunto objetivo se propone en *ingresos anuales por innovación* (mmusd) de **2500**, que es una puntuación de la desviación **-2500**. Este valor de los ingresos, corresponde a una puntuación de pertenencia de **0.05** con una posibilidad logarítimica natural de **-3.0** (ver **Tabla 6.3**). Por lo tanto, los casos con *ingresos anuales por innovación* (mmusd) de **2500** o inferiores (es decir, la desviación de las puntuaciones **-2500** o más bajas) se consideran totalmente fuera de la meta establecida, con puntuaciones de pertenencia **<= 0.05** y *posibilidad logaritmica natural* de pertenencia **<= 3.0.**
6. Una vez que estos tres valores (los dos umbrales y el punto de cruce) han sido seleccionados, es posible calibrar el grado de pertenencia en el conjunto objetivo. La tarea principal en este punto, es traducir los datos de *ingresos anuales por innovación* (mmusd) del punto de cruce (*columna 2*) en la métrica de la pertenencia por *posibilidad logarítmica natural*, con la utilización de los criterios externos que se han operacionalizado a través de las *tres referencias cualitativas* mencionadas anteriormente.
7. Para las puntuaciones de desviación por encima del *punto de cruce*, esta traducción se puede lograr, mediante la multiplicación de la puntuaciones de desviación correspondientes (*columna 2*, **Tabla 6.4**) por la relación de la pertenencia por *posibilidad logarítmica natural* con lo

marcado en la etiqueta para el *umbral de la plena pertenencia* (*threshold for full membership*, **3.0**, calculada en el *punto 4*) y a la puntuación de la desviación designado como el *umbral de plena pertenencia* (*threshold for full membership*, es decir, **20000 - 5000 = 15000**). Por lo que esta relación es **3 / 15000 o 0.0002**.

**8.** Para las puntuaciones de desviación por debajo del punto de cruce, esta traducción se puede logar, mediante la multiplicación de las puntuaciones de desviación correspondientes (*columna 3*, **Tabla 6.4**) por la relación de la pertenencia por *posibilidad logarítmica natural* con lo marcado en la etiqueta para el umbral de la no plena pertenencia (*threshold for full nob-membership*, **-3.0**, calculada en el *punto 5*) y a la puntuación de la desviación designado como el umbral de no plena pertenencia (*threshold for full non-membership*, es decir, **2500 - 5000 = - 2500**). Por lo que esta relación es **-3 / -2500, o 0.0012**.

**9.** Estos dos escalares, se muestran en la *columna 3,* y los productos de las *columnas 3 y 4* se muestran en la *columna 4*. Estos dos escalares, constituyen las pendientes de las dos líneas que se extienden desde el origen **(0.0)** a los dos puntos de umbral **(15000, 3)** y **( -2500, -3)** en la gráfica de las desviaciones del *ingresos anuales por innovación* (mmusd) en el punto de cruce (**eje X**) contra, la pertenencia por posibilidad logarítmica natural de pleno pertenencia (*full membership*) en el conjunto de las empresas innovadoras del clúster (**eje Y**). Por lo tanto, la *columna 4* muestra la traducción de la desviación de los *ingresos anuales por innovación* (mmusd) dentro de la métrica de pertenencia por posibilidad logarítmica natural, utilizando las tres referencias cualitativas para estructurar la transformación vía, los dos escalares. Los valores en la *columna 4,* en efecto, son los valores de ingresos anuales por innovación (mmusd) reflejando la pertenencia por *posibilidad logarítmica natural* en el conjunto de las empresas innovadoras del clúster, de una manera que estrictamente ajuste a los valores asignados a las tres referencias cualitativas (*umbral de plena pertenencia, el umbral de no plena pertenencia y el punto de cruce*). Por lo tanto, los valores en la *columna 4* no son puramente cambios de escala mecanicistas de los *ingresos anuales*

*por innovación* (mmusd) **I**, ya que reflejan la imposición de criterios externos cualitativos a través de las *tres referencias cualitativas*.

**10.** Es un pequeño paso de la pertenencia por posibilidad logarítmica natural, reportada en la *columna 4* para el grado de pertenencia. Solo es necesario aplicar la fórmula estándar para la conversión a *logaritmo natural* a puntajes con rangos de **0.0 a 1.0**, a saber:

$$\mathbf{(Y) = [exp\ (P)\ ] / [\ 1+ exp\ (P)]}$$

Donde:

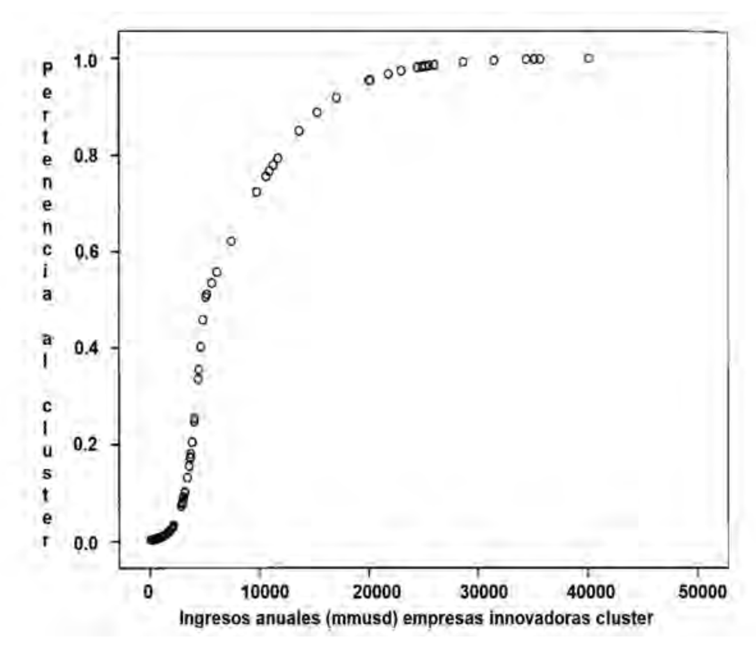
**exp.** Representa la exponenciación de la posibilidad del logaritmo natural para convertirlo a simples probabilidades. Estos pasos aparentemente complejos, son abreviados con el uso del software **fsQCA**.

**11.** Tenga en cuenta, que los valores de pertenencia reportados en la última columna de la **Tabla 6.4** se ajustan estrictamente a la distribución dictada por los tres referencias cualitativas mencionadas en el **punto 1**. Es decir, para nuestro ejemplo, el *umbral para la plena pertenencia* (*full-membership*) (**0.95**) está vinculado a un valor de *ingresos anuales por innovación* (mmusd) **I** de **20000**; el *punto de cruce* (**0.5**) está vinculado a un ingreso de **5000** y así sucesivamente.

**12.** Para más ilustración de los resultados del *método directo*, considere la **Figura 6.4**.

Que muestra un gráfico del grado de pertenencia al conjunto de empresas innovadoras del clúster vs. su ingresos anuales por innovación (mmusd), utilizando los datos de los **136** empresas incluídos en esta demostración. Como lo muestra el gráfico, la línea se aplana cuando se acerca **0.0** (*no pertenencia completa*) y **1.0** (*pertenencia completa*), consistente con la conceptualización del conjunto de grado de pertenencia. Lo que el gráfico no revela, es que la mayoría de las empresas innovadoras del clúster, se encuentran en la esquina inferior izquierda del gráfico, con un bajo ingresos anuales por innovación (mmusd) y la exclusión completa de la conjunto de empresas innovadoras (es decir, conjunto de puntuaciones de pertenencia < **0.05**).

**Figura 6.4. Gráfico de grado depertenencia en el conjunto de empresas competitivas de alta tecnología del clúster vs. ingresos anuales por innovación (mmusd). Método directo**




Fuente: Ragin (2008) con adaptación propia.

## Caso 2

Para ilustrar la importancia de los criterios de calibración externos, considere usar los mismos datos de *ingresos anuales por innovación* (mmusd) de la *columna 2* de la **Tabla 6.4** a calibrar grado de pertenencia del conjunto empresas competitivas de alta tecnología del clúster hipotético, que sean al menos moderadamente innovadoras. Debido a que la definición del conjunto objetivo ha sido cambiada, así también, debemos cambiar las tres referencias cualitativas. Las referencias apropiadas para el nuevo conjunto de empresas competitivas de alta tecnología del clúster, a calibrar de al menos moderadamente innovadores tienen un punto de cruce de **2500**; un

umbral de valor de pertenencia completa de **7500** y un umbral de valor no pertenencia completa de **1000**. la adecuada Los valores escalares asignados para este ejemplo, son **3 / 5000** para los casos arriba del valor de punto de cruce y **-3 / -1500** los casos debajo del valor de punto de cruce. Así, se establece la siguiente pregunta:

**P2.** ¿Cómo calibrar los datos de la **Tabla 6.5.** basados en sus ingresos anuales, por innovación **8 mmusd**) para que sean insumos de investigación, de *punto de cruce* = **2500**; superior al *punto de cruce*= **5000**; inferior al *punto de cruce*= **1500** consideradas moderadamente innovadoras;? El procedimiento completo se muestra en la **Tabla 6.5** utilizando los mismos casos como en la **Tabla 6.4**.

El punto clave de contraste entre la **Tabla 6.4 y 6.5** se muestra en la la *última columna*, las puntuaciones de pertenencia calibrados. Por ejemplo, con los *ingresos anuales por innovación* (mmusd) **(I)** de **2980**, la empresa competitiva de alta tecnología del clúster número **20** tiene una pertenencia de de **0.08** en el conjunto de los empresas innovadoras del clúster. Entre los miembros en el conjunto al menos moderadamente innovadores, es de **0.57**, lo que coloca que por encima del *punto de cruce*. Observe además, que en la **Tabla 6.5** se registran muchos más casos con puntuaciones de pertenencia cercanas a **1.0**, consistente con el simple hecho de que más empresas competitivas de alta tecnología del clúster tienen una alta calidad de pertenencia en el conjunto de empresas que son al menos moderadamente innovadoras en el conjunto de los empresas que son plenamente innovadoras. El contraste entre las **Tablas 6.4** y **6.5** subraya tanto la naturaleza del conocimiento dependiente de la calibración así como del impacto de la aplicación de diferentes estándares externos la misma medición (*ingresos anuales por innovación*, **I**). Una vez más, la clave para la comprensión de la calibración es de comprender la importancia de los criterios externos, que están basados, a su vez, en el conocimiento sustantivo empírico y de marco teórica, que los investigadores aportan a su investigación.

**Tabla 6.5. Caso hipotético: calibrando el grado de pertenencia de las empresas del clúster de alta tecnología del clúster, moderadamente innovadoras. Método directo**

<table>
<tr><th rowspan="6">Empresas competi-tivas de alta tecnología clúster</th><th colspan="7">Columnas</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th colspan="2">5</th><th>6</th></tr>
<tr><th>Ingresos anuales por innovación (mmusd)</th><th>Desvia-ciones del punto de cruce (crossover)</th><th>Escalares Puntos Cruce = 2500 Superior = 5000; Inferior= 1500</th><th>Posibilidad de pertenencia (odds of membership) por logaritmo natural</th><th colspan="2">Grado de pertenencia</th><th rowspan="4">Consistencia, Cobertura y Correlación</th></tr>
<tr><th rowspan="3">(I)</th><th rowspan="3">D = ( I – 2500 )</th><th>(X) 3 / 5000= 0.0006 3 / 1500= 0.0020</th><th rowspan="3">(P) = (D) * (X)</th><th colspan="2">(Y) = [exp (P) ] / [ 1+ exp (P)]</th></tr>
<tr><th rowspan="2">Σ ($X_i$) = 0.020</th><th>Y</th><th rowspan="2">Y Redondeo</th></tr>
<tr><th>Σ ($Y_i$) = 17.825</th></tr>
<tr><td>1</td><td>40110</td><td>37610</td><td>0.0006</td><td>22.566</td><td>1.00000</td><td>1.00</td><td rowspan="24">Consistencia<br>( $X_i$ < = $Y_i$ )<br>Σ [ min ( $X_i$ , $Y_i$ ) ] /<br>Dado que existen 4 valores = 1.0, estos se sustraen para precisar la *consistencia*, resultando (0.02-(4*0006)) / (0.02)= 0.88<br><br>***Cobertura***<br>( $X_i$ < = $Y_i$ )<br>Σ [ min ( $X_i$ , $Y_i$ ) ] / Σ ( $Y_i$ )<br>(0.02-(4*0006)) / (17.825)= 0.00098<br><br>**Factor de Correlación de Pearson**<br>-0.919</td></tr>
<tr><td>2</td><td>34400</td><td>31900</td><td>0.0006</td><td>19.14</td><td>1.00000</td><td>1.00</td></tr>
<tr><td>3</td><td>25200</td><td>22700</td><td>0.0006</td><td>13.62</td><td>1.00000</td><td>1.00</td></tr>
<tr><td>4</td><td>24920</td><td>22420</td><td>0.0006</td><td>13.452</td><td>1.00000</td><td>1.00</td></tr>
<tr><td>5</td><td>20060</td><td>17560</td><td>0.0006</td><td>10.536</td><td>0.99997</td><td>1.00</td></tr>
<tr><td>6</td><td>17090</td><td>14590</td><td>0.0006</td><td>8.754</td><td>0.99984</td><td>1.00</td></tr>
<tr><td>7</td><td>15320</td><td>12820</td><td>0.0006</td><td>7.692</td><td>0.99954</td><td>1.00</td></tr>
<tr><td>8</td><td>13680</td><td>11180</td><td>0.0006</td><td>6.708</td><td>0.99878</td><td>1.00</td></tr>
<tr><td>9</td><td>11720</td><td>9220</td><td>0.0006</td><td>5.532</td><td>0.99606</td><td>1.00</td></tr>
<tr><td>10</td><td>11290</td><td>8790</td><td>0.0006</td><td>5.274</td><td>0.99490</td><td>0.99</td></tr>
<tr><td>11</td><td>10940</td><td>8440</td><td>0.0006</td><td>5.064</td><td>0.99372</td><td>0.99</td></tr>
<tr><td>12</td><td>9800</td><td>7300</td><td>0.0006</td><td>4.38</td><td>0.98763</td><td>0.99</td></tr>
<tr><td>13</td><td>7470</td><td>4970</td><td>0.0006</td><td>2.982</td><td>0.95175</td><td>0.95</td></tr>
<tr><td>14</td><td>4670</td><td>2170</td><td>0.0006</td><td>1.302</td><td>0.78617</td><td>0.79</td></tr>
<tr><td>15</td><td>4100</td><td>1600</td><td>0.0006</td><td>0.96</td><td>0.72312</td><td>0.72</td></tr>
<tr><td>16</td><td>4070</td><td>1570</td><td>0.0006</td><td>0.942</td><td>0.71950</td><td>072</td></tr>
<tr><td>17</td><td>3740</td><td>1240</td><td>0.0006</td><td>0.744</td><td>0.67787</td><td>0.68</td></tr>
<tr><td>18</td><td>3690</td><td>1190</td><td>0.0006</td><td>0.714</td><td>0.67128</td><td>0.67</td></tr>
<tr><td>19</td><td>3590</td><td>1090</td><td>0.0006</td><td>0.654</td><td>0.65791</td><td>0.66</td></tr>
<tr><td>20</td><td>2980</td><td>480</td><td>0.0006</td><td>0.288</td><td>0.57151</td><td>0.57</td></tr>
<tr><td>21</td><td>1000</td><td>-1500</td><td>0.0020</td><td>-3</td><td>0.04743</td><td>0.05</td></tr>
<tr><td>22</td><td>650</td><td>-1850</td><td>0.0020</td><td>-3.7</td><td>0.02413</td><td>0.02</td></tr>
<tr><td>23</td><td>450</td><td>-2050</td><td>0.0020</td><td>-4.1</td><td>0.01630</td><td>0.02</td></tr>
<tr><td>24</td><td>110</td><td>-2390</td><td>0.0020</td><td>-4.78</td><td>0.00833</td><td>0.01</td></tr>
</table>

Notas: con propuestas de punto de cruce = 2500; superior al punto de cruce= 5000; inferior al punto de cruce= 1500; numerador 3, ver cálculo en los puntos 7,8 y 12 arriba mencionados.

Fuente: Elaboración propia.

## Método indirecto de calibración

A diferencia del *método directo*, que se basa en la especificación de los valores numéricos vinculados a tres referencias cualitativas arriba mencionadas, el *método indirecto* se basa en las amplias agrupaciones de casos, que proponga el investigador, según su grado de pertenencia en el conjunto objetivo. En esencia, el investigador realiza una clasificación inicial de casos en diferentes niveles de pertenencia para asignarlos como preliminares y posteriormente afina estos puntajes de pertenencia usando la escala de intervalos datos. Considere nuevamente los datos hipotéticos de las empresas innovadoras en clúster presentadas en la **Tabla 6.6**.

**Tabla 6.6. Caso hipotético: calibrando el grado de pertenencia de las empresas competitivas de alta tecnología cluster. Método indirecto**

| Empresas competitivas de alta tecnología clúster | Columnas | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| | Ingresos anuales Por innovación (mmusd) | Codificación cualitativa (asignado por el investigador empírico sustantivo) | Codificación cualitativa predicha (uso logit fraccional) | Resultados Tabla 5.2 Columna 6 |
| | I | C | P | Y |
| **1** | 40110 | 1.0 | 1.000 | 1.00 |
| **2** | 34400 | 1.0 | 1.000 | 1.00 |
| **3** | 25200 | 1.0 | 1.000 | 0.98 |
| **4** | 24920 | 1.0 | 1.000 | 0.98 |
| **5** | 20060 | 1.0 | 0.999 | 0.95 |
| **6** | 17090 | 0.8 | 0.991 | 0.92 |
| **7** | 15320 | 0.8 | 0.991 | 0.89 |
| **8** | 13680 | 0.8 | 0.977 | 0.85 |
| **9** | 11720 | 0.8 | 0.887 | 0.79 |
| **10** | 11290 | 0.8 | 0.868 | 0.78 |
| **11** | 10940 | 0.8 | 0.852 | 0.77 |
| **12** | 9800 | 0.6 | 0.793 | 0.72 |
| **13** | 7470 | 0.6 | 0.653 | 0.62 |
| **14** | 4670 | 0.4 | 0.495 | 0.40 |
| **15** | 4100 | 0.4 | 0.465 | 0.25 |
| **16** | 4070 | 0.4 | 0.463 | 0.25 |
| **17** | 3740 | 0.2 | 0.445 | 0.18 |
| **18** | 3690 | 0.2 | 0.442 | 0.17 |
| **19** | 3590 | 0.2 | 0.436 | 0.16 |
| **20** | 2980 | 0.2 | 0.397 | 0.08 |
| **21** | 1000 | 0.0 | 0.053 | 0.01 |
| **22** | 650 | 0.0 | 0.002 | 0.01 |
| **23** | 450 | 0.0 | 0.000 | 0.00 |
| **24** | 110 | 0.0 | 0.000 | 0.00 |

Fuente: Elaboración propia.

## Caso 3

**P3.** ¿Cómo calibrar los datos de la **Tabla 6.4.** basados en sus *ingresos anuales por innovación* (mmusd) **I**, para que sean insumos de investigación, de punto de cruce = **2500**; superior al punto de cruce = **5000**; inferior al punto de cruce = **1500**, con el método indirecto?

El *método indirecto* se realiza, como sigue:

1. El primer y más importante paso en el método indirecto, es que el investigador categorice los casos de manera cualitativa, según sus supuestos grados de pertenencia en el conjunto objetivo. Estas agrupaciones cualitativas deben ser preliminares y abiertas a revisión. Sin embargo, deben basarse lo más posible en el conocimiento de un marco teórico de la investigación así como a un referente empírico y sustantivo existente.
2. Las seis categorías cualitativas clave utilizadas en esta demostración se muestran en la **Tabla 6.7.**

**Tabla 6.7. Categorías cualitativas de 6 niveles**

| Item | Etiqueta | Puntuación de pertenencia |
|---|---|---|
| 1 | Dentro del objetivo establecido (*In the target*) | 1.0 |
| 2 | Principalmente pero no completamente en el objetivo establecido (*Mostly but not fully in the target set*) | 0.8 |
| 3 | Más dentro que fuera del objetivo establecido (*More in than out of the target set*) | 0.6 |
| 4 | Más afuera que en el objetivo establecido (*More out than in the target set*) | 0.4 |
| 5 | Principalmente pero no completamente fuera del objetivo establecido (*Mostly but not fully out of the target set*) | 0.2 |
| 6 | Fuera del objetivo establecido (*Out of the target*) | 0.0 |

Fuente: Ragin (2008) con adaptación propia.

Por supuesto, son posibles otros esquemas de codificación, utilizando tan solo tres categorías. El punto importante, es que la puntuación de estas categorías deben reflejar la estimación inicial del investigador del grado de pertenencia establecida de cada caso. Estas categorizaciones cualitativas, proporcionan la base para una calibración más afinadas y se muestran en la *columna 1* de la **Tabla 6.6**, usando valores numéricos explícitos, para reflejar estimaciones preliminares del grado de afiliación de conjunto. Los seis valores numéricos de la **Tabla 6.7** no son arbitrarios, por supuesto, pero se eligen como estimaciones aproximadas del grado de pertenencia específico a cada agrupación cualitativa.

**3.** El objetivo del método indirecto, es re-escalar el indicador de escala de intervalo para reflejar el conocimiento cualitativo de las agrupaciones de casos, clasificadas según el grado de pertenencia establecida. Estas interpretaciones cualitativas de casos, deben estar fundamentadas, inicialmente, en conocimiento empírico sustantivo. Cuanto más fuerte sea la base empírica para hacer las evaluaciones cualitativas de la pertenencia establecida, cuanto más preciso es la calibración de los valores del indicador de escala de intervalo como conjunto de puntuaciones a utilizar.

**4.** Tenga en cuenta, que las agrupaciones cualitativas implementadas en la **Tabla 6.6** han sido estructurados para que utilicen aproximadamente, los mismos criterios utilizados para estructurar las calibraciones que se muestran en la **Tabla 6.4.** Es decir, empresas innovadoras del clúster, con *ingresos anuales por innovación* (mmusd) **(I)** superior a **20000**, se han codificado como empresas totalmente innovadoras; *empresas competitivas de alta tecnología clúster* **5000** han sido codificados como más dentro que fuera y así. Al mantener la fidelidad de las referencias cualitativas utilizados en la **Tabla 6.4**, es posible comparar los resultados de los dos métodos. El método directo, utiliza especificaciones precisas de los puntos de referencia clave, mientras que el método indirecto requiere solo una amplia clasificación de casos.

**5.** El siguiente paso es usar las dos series reportadas en las *columnas 1 y 2* de la **Tabla 6.6** para estimar la codificación cualitativa predicha

de cada caso, utilizando el *ingresos anuales por innovación* (mmusd) (**I**) como la variable independiente y la codificaciones cualitativas (**C**) como la variable dependiente. La mejor técnica para realizar esta tarea, es aplicar un modelo logit fraccional (*polinomio fraccional*), mostrado en la *columna 3*, que se implementa en **STATA**. Los comandos a considerar *fracpoly glm qualcode intervv, family (binomial) lonk (logit)* y entonces *predict fzpred*, donde *qualcode* es la variable que implementa el investigador los seis valores codificados del conjunto de puntuaciones de pertenencia, mostrados en la **Tabla 6.5**; *intervv* es el nombre de la variable de intervalo de escala a utilizar para generar puntuaciones de pertenencia difusas; *fzpred* es el valor de predicción mostrando el resultado de las puntuaciones de pertenencia difusas.

**6.** Los valores predichos resultantes de este análisis, se presentan en la *columna 3* de la **Tabla 6.6.** Los valores reportados, se basan en un análisis que utiliza los **136** casos, no al subconjunto de **24** presentados en la tabla. Los valores predichos, en esencia, constituyen estimaciones de pertenencia difusa en el conjunto de empresas innovadoras del clúster basados en el *ingresos anuales por innovación (mmusd)* **I**, de la *columna 1* y el análisis cualitativo que produjo las codificaciones mostradas en la *columna 2*.

**7.** La comparación de los puntajes de pertenencia, establecidos en la *columna 5* de la **Tabla 6.4** (*método directo*) y la *columna 3* de la **Tabla 6.6** (*método indirecto*) se presenta en la *columna 4* de dicha **Tabla 6.6**. Esta revela grandes similitudes, pero también algunas diferencias importantes:

   **a.** Primero, observe que en la **Tabla 6.4** al implementar **20000** como el umbral de plena pertenencia en el conjunto de empresas innovadoras **(0.95).** En la **Tabla 6.6**, sin embargo, este valor de umbral cae muy por debajo del puntaje de la empresa innovadora **8 (13680).**

   **b.** Segundo, observe que el *método indirecto* revela un gran brecha que separa a la empresa competitiva de alta tecnología clúster **20 (0.397)** y el siguiente caso, la empresa competitiva de alta tecnología clúster

**21 (0.053).** Utilizando el *método directo*, sin embargo, esta brecha es mucho más estrecha, con la empresa competitiva de alta tecnología clúster **20** en **(0.08)** y la empresa competitiva de alta tecnología clúster **21** en **(0.01).** Estas diferencias, surgen a pesar del uso del mismo criterio general, se deducen indirectamente, del segundo método y su dependencia necesaria en la estimación de regresión. Todavía, si los investigadores carecieran de los criterios externos utilizados en el *método directo*, la comparación de las **Tablas 6.4** y **6.6** confirma que el *método indirecto* produce puntuaciones útiles de pertenencia.

## Cómo usar las mediciones calibradas

Las mediciones calibradas tienen muchos usos. Son especialmente útiles en evaluación de la teoría que se formula en términos de relaciones establecidas. Como se ha señalado en el capítulo 1, mientras que algunas teorías de las ciencias sociales aplican estrictamente la matemática, para su expresión, la gran mayoría lo hace de manera verbal, la cual a su vez, se formula casi por completo en términos de relaciones establecidas **(Ragin 2000, 2006b)**. Desafortunadamente, los científicos sociales han tardado en reconocer este hecho. Considere, por ejemplo, la declaración: realizan prácticas de innovación del modelo de negocios, las empresas competitivas de alta tecnología clúster, como en muchas declaraciones de este tipo, la afirmación es esencialmente, que las instancias del conjunto realizan prácticas de innovación del modelo de negocios constituyen un subconjunto de instancias, del conjunto mencionado segundo (*empresas competitivas de alta tecnología clúster*). Es muy recomendable, al plantear los argumentos, que primero se mencione el subconjunto (*resultado*) y después el conjunto (*causa*) para facilitar los planteamientos. Un examen detallado de la mayoría de las teorías de las ciencias sociales **(Ragin, 2000)**, revela que la gran mayoría de los argumentos están compuestos en gran parte, de declaraciones que

describen primeramente, conjuntos de relaciones establecidas, que subconjunto de datos.

Estos conjuntos de relaciones establecidas, a su vez, pueden involucrar una variedad de diferentes tipos de conexiones empíricas: *descriptivas, constitutivas, o causal, entre otros.* Los conjuntos de relaciones establecidas, por ejemplo, con los realizan prácticas de innovación del modelo de negocios como un subconjunto de empresas competitivas de alta tecnología clúster, descritos anteriormente, también es compatible con un tipo de argumento causal, a saber, como que el empresas competitivas de alta tecnología clúster es suficiente pero no es necesario para realizar prácticas de innovación del modelo de negocios. En argumentos de este tipo, si la causa (empresas competitivas de alta tecnología clúster) está presente, entonces el resultado (realizar prácticas de innovación del modelo de negocios) también debería estar presente. Sin embargo, en casos donde del resultado (realizar prácticas de innovación del modelo de negocios) sin la causa (empresas competitivas de alta tecnología clúster) no cuenta ni socava el argumento empresas competitivas de alta tecnología clúster de que sea suficiente para realizar prácticas de innovación del modelo de negocios (aunque tales casos, socavarían dramáticamente la correlación). Es decir, estas instancias del resultado sin la causa, se deben a la existencia de rutas alternativas para estos resultados (por ejemplo, la imposición de una práctica especial de innovación por parte de una empresa competitiva de baja tecnología clúster con ciertas características). Así, en situaciones donde las instancias de una *condición causal* constituyen un subconjunto de instancias del resultado, un investigador puede afirmar que la causa es suficiente pero no necesaria, para el resultado. Como siempre, las afirmaciones de este tipo no pueden basarse simplemente en la demostración de la relación del subconjunto de datos. Los investigadores deberán reunir tanta evidencia corroborante como sea posible, cuando se hace cualquier tipo de afirmación causal.

## Consistencia

Antes del advenimiento de los conjuntos de *datos difusos* **(Zadeh, 1965, 1972, 2002; Lakoff, 1973)**, muchos científicos sociales desdeñaron el análisis de las relaciones dentro de los marcos teóricos porque tales análisis requieren el uso de variables de escala categórica (es decir, conjuntos binarios) que, a su vez, a menudo requiere de la *dicotomización* de *escalas de intervalo y de razón*. Por ejemplo, utilizando conjuntos de datos nítidos para evaluar una declaración proveniente del marco teórico sobre empresas competitivas de alta tecnología clúster, se podría requerir a un investigador las clasifique dos grupos: innovadoras y no innovadores, utilizando su prácticas de innovación del modelo de negocios. Tales prácticas, a menudo son criticadas porque los investigadores pueden manipular puntos clave de información, cuando se dicotomizan las variables de escala de intervalo y de razón, en formas que mejoren la *consistencia* de la evidencia del marco teórico de la investigación. Sin embargo, como se demuestra aquí, es posible calibrar el grado de pertenencia en conjuntos de relaciones y así evitar dicotomizaciones arbitrarias. Como se ha mencionado, el subconjunto de *datos difusos,* se establece demostrando que las puntuaciones de pertenencia en un conjunto de datos, son consistentemente menores que o igual a los puntajes de pertenencia en otro. En otras palabras, si para cada caso el grado de pertenencia en el conjunto **X** menor que o igual a, hasta al grado de pertenencia en el conjunto **Y**, entonces el conjunto **X** es un subconjunto del conjunto **Y**. Por supuesto, los datos de ciencias sociales rara vez son perfectos, y se debe tener en cuenta las acciones de corrección. Es posible evaluar el grado de *consistencia* de la evidencia empírica, con subconjunto de datos utilizando el fórmula simple descrita:

$$\textbf{Consistencia} \; ( X_i <= Y_i ) = \Sigma \, [ \, \min ( X_i , Y_i ) \, ] \, / \, \Sigma \, (X_i)$$

Donde:

$X_i$ **.** Grado de pertenencia en el conjunto de datos de **X.**

$Y_i$ **.** Grado de pertenencia en el conjunto de datos de **Y**

**(** $X_i <= Y_i$ **).** Es el subconjunto de datos en cuestión.

**min.** Selección de más bajo de las dos puntuaciones o *scores*.

## Caso 4

**P4.** ¿Cómo calibrar y lograr la *consistencia* de los datos de la **Tabla 6.8**. basados en la práctica de la innovación del modelo de negocios**,** para que sean insumos de investigación, de punto de cruce = **2**; superior al punto de cruce = **7**; inferior al punto de cruce = **-5**?

Por ejemplo, considere la *consistencia* de la evidencia empírica, con la afirmación de que el conjunto de prácticas de innovación del modelo de negocios (vea cómo se calibró en la **Tabla 6.4**) constituye un subconjunto del conjunto empresas competitivas de alta tecnología clúster, utilizando datos sobre **24/136** empresas. La calibración de la pertenencia, en el conjunto de empresas competitivas de alta tecnología clúster, utilizando el método directo, se muestra en la **Tabla 6.8**.

**Tabla 6.8. Caso hipotético: calibrando el grado de pertenencia en el conjunto de empresas de alta tecnología clúster. Método directo**

| Columna | | | | | | |
|---|---|---|---|---|---|---|
| Empresas competitivas de alta tecnología clúster | 1 | 2 | 3 | 4 | 5 | 6 |
| | Práctica de innovación del modelo de negocios | Desviaciones del punto de cruce (crossover) | Escalares Punto cruce = 2 superior = 7 Inferior= -5 | Posibilidad de pertenencia (odds of membership) por logaritmo natural | Grado de pertenencia | Consistencia , Cobertura y Correlación |
| | I | D= ( I – 2) | X<br>3 / 7= 0.430<br>3 / 5= 0.600<br>$\Sigma ( X_i ) =$ 12.19 | (P) = (D) * (X) | (Y) = [exp (P) ] / [ 1+ exp (P)]<br>$\Sigma = ( Y_i )$ 12.09 | En referencia a los puntos tratados en este apartado, se tiene:<br>• Punto 5: **Consistencia** $( X_i <= Y_i ) = \Sigma [ \min ( X_i , Y_i ) ] / \Sigma ( X_i )$<br>Esto es: 12.09/ 12.19= 0.0991<br>• Punto 8 **Cobertura** $( X_i <= Y_i ) = \Sigma [ \min ( X_i , Y_i ) ] / \Sigma (Y_i )$<br>Esto es: 12.09/12.09= 1<br>• Punto 9: **Correlación de Pearson** -0.814 |
| **1** | 10 | 8.00 | 0.430 | 3.44 | 0.97 | |
| **2** | 10 | 8.00 | 0.430 | 3.44 | 0.95 | |
| **3** | 9 | 7.00 | 0.430 | 3.01 | 0.93 | |
| **4** | 8 | 6.00 | 0.430 | 2.58 | 0.90 | |
| **5** | 7 | 5.00 | 0.430 | 2.15 | 0.90 | |
| **6** | 7 | 5.00 | 0.430 | 2.15 | 0.85 | |
| **7** | 6 | 4.00 | 0.430 | 1.72 | 0.85 | |
| **8** | 6 | 4.00 | 0.430 | 1.72 | 0.78 | |
| **9** | 5 | 3.00 | 0.430 | 1.29 | 0.78 | |
| **10** | 5 | 3.00 | 0.430 | 1.29 | 0.70 | |
| **11** | 4 | 2.00 | 0.430 | 0.86 | 0.61 | |
| **12** | 3 | 1.00 | 0.430 | 0.43 | 0.50 | |
| **13** | 2 | 0.00 | 0.430 | 0.00 | 0.50 | |
| **14** | 2 | 0.00 | 0.600 | 0.00 | 0.35 | |
| **15** | 1 | -1.00 | 0.600 | -0.60 | 0.23 | |
| **16** | 0 | -2.00 | 0.600 | -1.20 | 0.14 | |
| **17** | -1 | -3.00 | 0.600 | -1.80 | 0.08 | |
| **18** | -2 | -4.00 | 0.600 | -2.40 | 0.05 | |
| **19** | -3 | -5.00 | 0.600 | -3.00 | 0.03 | |
| **20** | -4 | -6.00 | 0.600 | -3.60 | 0.01 | |
| **21** | -5 | -7.00 | 0.600 | -4.20 | 0.01 | |
| **22** | -6 | -8.00 | 0.600 | -4.80 | 0.00 | |
| **23** | -7 | -9.00 | 0.600 | -5.40 | 0.00 | |
| **24** | -8 | -10.00 | 0.600 | -6.00 | 0.97 | |

Notas: con propuestas de punto de cruce = 2; superior al punto de cruce= 7; inferior al punto de cruce= 5; numerador 3, ver cálculo en los puntos 7,8 y 12 de método directo de calibración.

Fuente: Elaboración propia.

Para esta demostración, se toman en cuenta:

1. La variable práctica de innovación del modelo de negocios, la cual, tiene un rango de calificación entre **–10 y +10.** Existe la posibilidad considerar variables, que incluso sean deficientes, pero que son utilizadas dada su popularidad. Esta práctica es común en las ciencias sociales. Se presentan puntuaciones para **24** de los **136** países, incluidos en la calibración en la *columna 1* de la **Tabla 6.8**. Estos casos específicos, fueron seleccionados para proporcionar una gama de dichas puntuaciones a considerar.
2. La *columna 2*, muestra las desviaciones desde el punto de cruce con una puntuación de práctica de actividades de innovación del modelo de negocios de 2.
3. La *columna 3* muestra los escalares utilizados para transformar los puntajes de desviación de la práctica de actividades de innovación del modelo de negocios, en la métrica de pertenencia por probabilidades logarítmica natural en el conjunto de las empresas competitivas de alta tecnología clúster, de tal manera que:
   - **a.** El umbral de pertenencia completa (*full membership*) en el conjunto de las empresas competitivas de alta tecnología clúster se considera con práctica de actividades de innovación del modelo de negocios, con un puntaje de **9**.
   - **b.** Los casos de producto escalar como punto de cruce tienen una puntuación de **2.**
   - **c.** Los casos de producto escalar por arriba del punto de cruce con una puntuación de **3 / 7.**
   - **d.** El umbral de no pertenencia completa (*full non-membership*) en el conjunto de las empresas competitivas de alta tecnología clúster se considera con práctica de actividades de innovación del modelo de negocios, con un puntaje de **-3.**
   - **e.** Los casos de producto escalar por abajo del punto de cruce con un puntaje de **3 / -5.**

4. La *columna 3*, muestra el producto de los puntajes de desviación y los escalares mientras la *columna 4*, reporta los puntajes de pertenencia calibrados, usando el procedimiento descrito en la **Tabla 6.4.**
5. Aplicando la fórmula para la *consistencia* del marco teórico de la investigación:

$$\textbf{Consistencia} \ (X_i <= Y_i) = \Sigma \ [\min (X_i, Y_i)] \ / \ \Sigma \ (X_i)$$

   Con el argumento de que el conjunto de prácticas de innovación del modelo de negocios constituye un subconjunto de los empresas competitivas de alta tecnología clúster, es de **0.99**. (**1.0** indica *consistencia* perfecta).
6. Igualmente, la *consistencia* de la evidencia con el argumento de que el conjunto de al menos moderadamente innovadores (como se calibra en la **Tabla 6.5**) constituye un subconjunto del conjunto de empresas competitivas de alta tecnología clúster es **0.88**.
7. En resumen, ambos subconjuntos las relaciones son altamente consistentes, proporcionando un amplio soporte para ambas declaraciones (las empresas competitivas de alta tecnología clúster practican innovación en el modelo de negocios y las empresas que son al menos, moderadamente innovadoras son empresas competitivas de alta tecnología clúster). Asimismo, ambos análisis apoyan el argumento de que las prácticas de innovación del modelo de negocios son suficientes pero no necesarias para ser empresas competitivas de alta tecnología clúster. Tenga en cuenta, sin embargo, que del conjunto de al menos moderadamente innovadoras es un conjunto más inclusivo, con alto puntaje de pertenencia promedio. Mientras más grande sea el puntaje de pertenencia de la *condición causal* más difícil es satisfacer la igualdad del subconjunto de datos **(** $X_i <= Y_i$ **).** Las dos formulaciones también difieren sustancialmente en la *cobertura* de su marco teórico de investigación.
8. La *cobertura*:

$$\textbf{Cobertura} = (X_i <= Y_i) = \Sigma \ [\min (X_i, Y_i)] \ / \ \Sigma \ (Y_i)$$

   es un indicador de importancia o peso empírico **(Ragin, 2006b)**. Muestra la proporción de la suma de los resultados de los puntajes de pertenencia

(en nuestro ejemplo, el conjunto de empresas competitivas de alta tecnología cluster) que está cubierto por una *condición causal*. Es de observar el alto contraste que guardan las situaciones de *cobertura*:

**a.** De empresas competitivas de alta tecnología cluster con prácticas de innovación en el modelo de negocios es de **1.00** (**Tabla 6.8**).

**b.** De empresas competitivas de alta tecnología cluster con ingresos de innovación moderadamente innovadoras de **0.00** (**Tabla 6.5**).

Estos resultados invitan a una reconsideración, más que nada, partiendo de los argumentos empíricos y sustantivos y posteriormente, en el marco teórico de la investigación de ambos casos.

**9.** El factor de *correlación de Pearson*:

**a.** De empresas competitivas de alta tecnología clúster con prácticas de innovación en el modelo de negocios es de **-0.814** (**Tabla 6.8**).

**b.** Entre empresas competitivas de alta tecnología clúster con ingresos de innovación moderadamente innovadoras es de **-0.919** (**Tabla 6.5**).

Estos resultados invitan a una reconsideración, más que nada, partiendo de los argumentos del marco teórico, y posteriormente, a los empíricos y sustantivos de ambos casos.

**10.** Por lo tanto, utilizando métodos de análisis del marco teórico, es posible demostrar que la pertenencia en el conjunto de empresas competitivas de alta tecnología clúster con prácticas de innovación en el modelo de negocios, es más que suficiente; así, no se requiere plena pertenencia del conjunto de empresas competitivas de alta tecnología clúster con ingresos de innovación moderadamente innovadoras, aunque habrá que revisar los marcos teóricos de la investigación correspondientes ya que en ambas tablas, las correlaciones presentan un sentido inverso.

Como se explicó en el capítulo 1, es muy difícil evaluar los argumentos del marco teórico de la investigación, utilizando métodos correlacionales. Las tres fuentes principales de esta dificultad, son **(Ragin, 2008)**:

**1.** Las afirmaciones del marco teórico de la investigación, son sobre los tipos de casos; correlaciones que conciernen a las relaciones entre varia-

bles. La declaración: realizan prácticas de innovación del modelo de negocios, las empresas competitivas de alta tecnología clúster (es decir, realizan prácticas de innovación constituyen un subconjunto de empresas competitivas) invoca casos, no dimensiones de variación de cruce entre empresas. Este enfoque en casos, como instancias de conceptos, sigue directamente de la naturaleza del marco teórico de la investigación en las ciencias sociales. El cálculo de una correlación, por el contrario, se basa en un interés en evaluar qué tan bien las dimensiones de la variación entre sí a partir de una muestra o población y no se encuentra interesada en un conjunto de casos *per se*. En el estudio de casos, un conjunto de datos podría no incluir una sola empresa competitiva de alta tecnología clúster o una sola práctica de innovación en el modelo de negocios, pero en el estudio correlacional, el investigador aún podría calcular una correlación entre el grado de empresa competitiva de alta tecnología clúster y el grado de práctica de innovación del modelo de negocios.

**2.** Los argumentos correlacionales son completamente simétricos, mientras que se basan en el marco teórico de la investigación, son casi siempre asimétricos. La correlación entre la práctica de innovación del modelo de negocios y empresas competitivas de alta tecnología clúster (tratando a ambos como variables convencionales) se debilita por el hecho de que las prácticas de innovación del modelo de negocios se realizan también, por empresas competitivas de diferentes niveles de tecnología clúster. Sin embargo, tales casos no desafían al marco teórico de la investigación fortaleciendo o debilitando su *consistencia*. El argumento teórico en cuestión, aborda las cualidades de las empresas competitivas de alta tecnología: que practican innovación en el modelo de negocios, y no hace afirmaciones específicas sobre diferencias relativas entre empresas competitivas de alta tecnología clúster de menor o mayor nivel en su grado de prácticas de innovación en el modelo de negocios. Nuevamente, el análisis del marco téorico es fiel a la comunciación y formulación verbal, los cuales son típicamente asimétricas; la correlación no lo es.

3. Las correlaciones, son insensibles a las calibraciones implementadas por los investigadores. El contraste entre las **Tablas 6.4 y 6.5** son significativas, desde un punto de vista del marco teórico de la investigación establecida. El conjunto representado en la **Tabla 6.5** es más inclusivo y por lo tanto, proporciona un marco teórico de la investigación como prueba más exigente de la conexión entre prácticas de innovación del modelo de negocios y las empresas competitivas de alta tecnología clúster. A partir de una perspectiva correlacional, sin embargo, se considera que hay poca diferencia entre los dos factores. De hecho, el factor de *correlación de Pearson* de las **Tablas 6.5 y 6.8** aunque en magnitud, con diferencias no relevantes pero con un sentido inverso sugieren una revisión del marco teórico de lai investigación. La insensibilidad de correlación con la calibración se deduce directamente del hecho de que la correlación es dependiente de las desviaciones en las mediciones específicas de la muestra de tendencia central basadas en la media. Por esta razón, la correlación es incapaz de analizar las relaciones del marco teórico de la investigación y, en consecuencia, no puede usarse para *evaluar con la necesidad y suficiencia* de los causales **(Ragin, 2008)**.

Este capítulo demuestra la utilidad del conjunto de *datos difusos* así como su calibración. Es importante tener la capacidad de evaluar, no solo el más vs. menos (*medición no calibrada*) sino también mucho vs. poco (*medición calibrada*). El uso de mediciones calibradas fundamenta las ciencias sociales en el conocimiento empírico sustantivo y mejora la relevancia de los resultados de la investigación social, para la práctica y las cuestiones de normatividad de uso.

Los conjuntos de datos difusos, son especialmente potentes como portadores de calibración. Ofrecen herramientas de medición que trascienden lo cuantitativo y cualitativo en las ciencias sociales, especialmente orientados a la administración de la innovación.

Las prácticas actuales en ciencias sociales cuantitativas, socavan gravemente la atención por la calibración en los investigadores **(Ragin, 2008)**. Estas dificultades, se derivan de la confianza en el enfoque basado en el modelo de indicadores para la medición, que requiere solo variación a través de los factores o variables de la muestra y trata, todas las variaciones como igualmente significativas.

Las limitaciones de dicho enfoque, se refuerzan por las técnica correlacionales, que son insensibles a las calibraciones implementadas por investigadores. La confianza en las desviaciones de la media tienden a neutralizar el impacto de cualquier calibración directa implementada por el investigador. Una dificultad adicional, surge cuando es reconocido que casi toda la teoría de las ciencias sociales, se establece básicamente a partir del marco teórico de la investigación y que los métodos correlacionales son incapaces de evaluar. La naturaleza del marco teórico, generalmente no es reconocido por los investigadores de las ciencias sociales, ni el hecho de que la evaluación de argumentos provenientes de dicho marco y su calibración vayan de la mano. El análisis del marco teórico de la investigación, sin una cuidadosa calibración de la pertenencia del conjunto de datos, es un ejercicio inútil. Se deduce, que los investigadores deban de ser fieles a sus teorías, identificando claramente, los conjuntos de relaciones objetivo, que corresponden a los conceptos centrales de sus teorías y especificando los criterios externos útiles que pueden usarse para guiar la calibración de la pertenencia establecida.

## Resumen fsQCA para la calibración de datos difusos (*fuzzy-sets*). Método directo

A continuación, lo listamos en la **Tabla 6.9.**

**Tabla 6.9. Resumen fsQCA para la calibración de datos difusos (*fuzzy-sets*). Metodo directo**

| Item | Etapa | Descripción |
|---|---|---|
| 1 | **Creando el conjunto de datos** | Esta etapa se puede lograr usando **fsQCA:**<br>**a.** Que incluye procedimientos para ingresar datos directamente y para importar conjuntos de datos de otros programas (por ejemplo, archivos delimitados por comas desde Excel o archivos delimitados por tabulaciones desde **SPSS**). Crea y/o recupera su conjunto de datos. Por ejemplo, puede tener un archivo **SPSS** o **Excel**, con los datos de escala de intervalo o razón, relevantes. Guarda estos archivos como delimitados por comas o delimitados por tabulaciones con nombres de variables simples en la primera fila del archivo.<br>**b.** Los archivos importados deben tener nombres de variables simples (sin espacios incrustados o puntuación) como la primera línea del conjunto de datos.<br>**c.** Los datos faltantes deben ingresarse como espacios en blanco. Asegúrese de que los datos faltantes, estén en blanco y no se encuentren asignados a un código especial (por ejemplo, –999). |
| 2 | **Ajustes** | Con sus datos en la ventana de hoja de cálculo de datos de **fsQCA**, haga clic en el menú Variables (*Variables*); posteriormente haga clic en Calcular (*Compute*). |
| 3 | **Asignando nombre al conjunto difuso** | En el cuadro de diálogo de Calcular (*Compute*), asigne un nombre al conjunto difuso (*fuzzy-sets*) de destino. Seleccione un nombre simple (de dos a ocho caracteres), utilizando caracteres alfanuméricos estándar caracteres y sin espacios, guiones o puntuación. |
| 4 | **Calibración** | Para lograrlo:<br>**a.** Haga clic en Calibrar (*Calibrate* (x, n1, n2, n3)) en el menú Funciones (*Functions*) y posteriormente haga clic en la flecha hacia arriba que está al lado de la palabra *Functions*.<br>**b.** A continuación, Calibrar (*Calibrate (,,,)*) aparecerá en el campo Expresión (*Expression*) del cuadro de diálogo. |

| | | |
|---|---|---|
| **5** | **Ajuste de la calibración** | Para lograrlo:<br>**a.** Edite la expresión para que Calibrar (*Calibrate* (,,,)) y se convierta en algo así como *Calibrate (intvar, 25,10,2)*,<br>**b.** *intvar* es el nombre de la variable de intervalo o razón existente en el archivo,<br>**c.** El primer número es el valor de *intvar* que ha elegido como umbral para la pertenencia completa (*full membership*) en el conjunto objetivo (puntaje difuso = **0.95**),<br>**d.** El segundo número, es el valor de *intvar* que ha seleccionado para el punto de cruce (*crossover*) (puntaje difuso = **0.5**), y<br>**e.** El tercer número, es el valor de *intvar* que ha seleccionado para el umbral para de no pertenencia completa (*full non-membership*) en el conjunto objetivo (puntaje difuso = **0.05**). |
| **6** | **Verificación** | Para lograrlos:<br>**a.** Haga clic en Aceptar (*OK*). Verifique la hoja de cálculo de datos para asegurarse de que salió como esperaba.<br>**b.** Es posible ordenar la variable original de escala de intervalo en descendente o ascendente utilizando los menús desplegables.<br>**c.** Haga clic en cualquier caso en la columna que desea ordenar, posteriormente haga clic en Casos (*Cases*) y entonces Ordenar (*Sort*) (ascendente o descendente).<br>**d.** Posteriormente, puede verificar las correspondientes puntuaciones difusas para ver si se ajustan a su variable de escala de intervalo o de razón, de la manera que pretendía. |

Fuente: Ragin (2008) con adaptación propia.

## CAPÍTULO 7.
# Pensamiento configuracional en fsQCA

Este capítulo esboza las características principales del pensamiento *configuracional*, que es la base de muchas formas de análisis en las ciencias sociales. Los investigadores cualitativos, tienden especialmente a pensar en términos de combinaciones y *configuraciones* debido a su interés en el contexto y en la comprensión fenómenos sociales de manera integral. En la investigación orientada a casos, los investigadores cualitativos a menudo piensan en las *condiciones causales* en términos de lo que podría llamarse trayectorias causales, es decir, las condiciones causalmente relevantes que se combinan para producir un resultado dado. Este interés en las combinaciones de las causas, encaja perfectamente con el cómo suceden las cosas. Pensar en términos de trayectorias causales es pensar holísticamente y comprender dichas condiciones causalmente relevantes como intersecciones de fuerzas y eventos.

Este capítulo primero, examina el pensamiento *configuracional* para el estudio de casos y posteriormente, muestra cómo la investigación puede basarse en el estudio de un solo caso, utilizando la idea de trayectorias causales. Después de discutir brevemente los problemas que el estudio de trayectorias causales exhibe para el enfoque cuantitativo de investigación convencional, el capítulo se convierte en un tema clave en el pensamiento *configuracional*, utilizando los conjuntos de *datos difusos*, en cómo evaluar el grado de pertenencia de un caso en una configuración, lo cual es la base para la evaluación de las trayectorias causales. Además, el capítulo muestra cómo usar las medidas de *consistencia* y *cobertura* teóricas del marco teórico de la investigación presentados, para comparar las trayectorias causales y evaluar su poder explicativo.

## Relación con el punto de vista cualitativo

La centralidad del pensamiento *configuracional* para la investigación social, más relevante, es en una de sus formas más elementales: el estudio de caso. Suponga, por ejemplo, que un investigador argumenta que México se posiciona, de acuerdo con el *Global Innovation Index Report* **(GII, 2019)** desglosados en la **Tabla 7.1, 7.2 , 7.3 y 7.4** como componentes del *subíndice de innovación de entrada* (*innovation input sub-index*), que presentan puntuación y posición.

**Tabla 7.1. Puntuación y posición de México de acuerdo con el factor subíndice de innovación de entrada (*innovation input sub-índex)***

| Item | Variables | Puntuación | Posición |
|---|---|---|---|
| 1 | Instituciones (*Institutions*) | 62.8 | 66 |
| 2 | Capital humano e investigación (*Human capital & Research*) | 33.4 | 54 |
| 3 | Infraestructura (*Infrastructure*) | 48.3 | 59 |
| 4 | Sofisticación de mercado (*Market sophistication*) | 49.9 | 57 |
| 5 | Sofisticación de negocio (*Business sophistication*) | 29.4 | 73 |

Fuente: GII (2019).

Con valores intermedios, siendo especialmente bajos, como la sofisticación de negocio y capital humano e investigación, en la tabla de clasificación.

Los componentes del subíndice de innovación de salida (*innovation output sub-index*), compuesto por:

**Tabla 7.2. Puntuación y posición de México de acuerdo con el factor subíndice de innovación de salida (*innovation output sub-índex)***

| Item | Variables | Puntuación | Posición |
|---|---|---|---|
| 6 | Conocimiento y tecnología producidas (*Knowledge and technology outputs*) | 25.5 | 50 |
| 7 | Creatividad producida (*Creative outputs*) | 29.2 | 55 |

Fuente: GII (2019).

Y que dadas las entradas, los productos son bajos.

**Tabla 7.3. Posicion de México de acuerdo al *Global Innovation Index* 2019 nivel general**

| Country/Economy | Score (0–100) | Rank | Income | Rank | Region | Rank | Median 33.86 |
|---|---|---|---|---|---|---|---|
| Switzerland | 67.24 | 1 | HI | 1 | EUR | 1 | |
| Sweden | 63.65 | 2 | HI | 2 | EUR | 2 | |
| United States of America | 61.73 | 3 | HI | 3 | NAC | 1 | |
| Netherlands | 61.44 | 4 | HI | 4 | EUR | 3 | |
| United Kingdom | 61.30 | 5 | HI | 5 | EUR | 4 | |
| Finland | 59.83 | 6 | HI | 6 | EUR | 5 | |
| Denmark | 58.44 | 7 | HI | 7 | EUR | 6 | |
| Singapore | 58.37 | 8 | HI | 8 | SEAO | 1 | |
| Germany | 58.19 | 9 | HI | 9 | EUR | 7 | |
| Israel | 57.43 | 10 | HI | 10 | NAWA | 1 | |
| Republic of Korea | 56.55 | 11 | HI | 11 | SEAO | 2 | |
| Ireland | 56.10 | 12 | HI | 12 | EUR | 8 | |
| Hong Kong, China | 55.54 | 13 | HI | 13 | SEAO | 3 | |
| China | 54.82 | 14 | UM | 1 | SEAO | 4 | |
| Japan | 54.68 | 15 | HI | 14 | SEAO | 5 | |
| France | 54.25 | 16 | HI | 15 | EUR | 9 | |
| Canada | 53.88 | 17 | HI | 16 | NAC | 2 | |
| Luxembourg | 53.47 | 18 | HI | 17 | EUR | 10 | |
| Norway | 51.87 | 19 | HI | 18 | EUR | 11 | |
| Iceland | 51.53 | 20 | HI | 19 | EUR | 12 | |
| Austria | 50.94 | 21 | HI | 20 | EUR | 13 | |
| Australia | 50.34 | 22 | HI | 21 | SEAO | 6 | |
| Belgium | 50.18 | 23 | HI | 22 | EUR | 14 | |
| Estonia | 49.97 | 24 | HI | 23 | EUR | 15 | |
| New Zealand | 49.55 | 25 | HI | 24 | SEAO | 7 | |
| Czech Republic | 49.43 | 26 | HI | 25 | EUR | 16 | |
| Malta | 49.01 | 27 | HI | 26 | EUR | 17 | |
| Cyprus | 48.34 | 28 | HI | 27 | NAWA | 2 | |
| Spain | 47.85 | 29 | HI | 28 | EUR | 18 | |
| Italy | 46.30 | 30 | HI | 29 | EUR | 19 | |
| Slovenia | 45.25 | 31 | HI | 30 | EUR | 20 | |
| Portugal | 44.65 | 32 | HI | 31 | EUR | 21 | |
| Hungary | 44.51 | 33 | HI | 32 | EUR | 22 | |
| Latvia | 43.23 | 34 | HI | 33 | EUR | 23 | |
| Malaysia | 42.68 | 35 | UM | 2 | SEAO | 8 | |
| United Arab Emirates | 42.17 | 36 | HI | 34 | NAWA | 3 | |
| Slovakia | 42.05 | 37 | HI | 35 | EUR | 24 | |
| Lithuania | 41.46 | 38 | HI | 36 | EUR | 25 | |
| Poland | 41.31 | 39 | HI | 37 | EUR | 26 | |

*Continuación Tabla 7.3.*

| Country/Economy | Score (0–100) | Rank | Income | Rank | Region | Rank | Median 33.86 |
|---|---|---|---|---|---|---|---|
| Bulgaria | 40.35 | 40 | UM | 3 | EUR | 27 | |
| Greece | 38.90 | 41 | HI | 38 | EUR | 28 | |
| Viet Nam | 38.84 | 42 | LM | 1 | SEAO | 9 | |
| Thailand | 38.63 | 43 | UM | 4 | SEAO | 10 | |
| Croatia | 37.82 | 44 | HI | 39 | EUR | 29 | |
| Montenegro | 37.70 | 45 | UM | 5 | EUR | 30 | |
| Russian Federation | 37.62 | 46 | UM | 6 | EUR | 31 | |
| Ukraine | 37.40 | 47 | LM | 2 | EUR | 32 | |
| Georgia | 36.98 | 48 | LM | 3 | NAWA | 4 | |
| Turkey | 36.95 | 49 | UM | 7 | NAWA | 5 | |
| Romania | 36.76 | 50 | UM | 8 | EUR | 33 | |
| Chile | 36.64 | 51 | HI | 40 | LCN | 1 | |
| India | 36.58 | 52 | LM | 4 | CSA | 1 | |
| Mongolia | 36.29 | 53 | LM | 5 | SEAO | 11 | |
| Philippines | 36.18 | 54 | LM | 6 | SEAO | 12 | |
| Costa Rica | 36.13 | 55 | UM | 9 | LCN | 2 | |
| Mexico | 36.06 | 56 | UM | 10 | LCN | 3 | |
| Serbia | 35.71 | 57 | UM | 11 | EUR | 34 | |
| Republic of Moldova | 35.52 | 58 | LM | 7 | EUR | 35 | |
| North Macedonia | 35.29 | 59 | UM | 12 | EUR | 36 | |
| Kuwait | 34.55 | 60 | HI | 41 | NAWA | 6 | |
| Iran (Islamic Republic of) | 34.43 | 61 | UM | 13 | CSA | 2 | |
| Uruguay | 34.32 | 62 | HI | 42 | LCN | 4 | |
| South Africa | 34.04 | 63 | UM | 14 | SSF | 1 | |
| Armenia | 33.98 | 64 | UM | 15 | NAWA | 7 | |
| Qatar | 33.86 | 65 | HI | 43 | NAWA | 8 | |

Fuente: GII (2019).

## Tabla 7.4. Posición de México de acuerdo con el Global Innovation Index 2019 nivel detallado por factor

| Output rank | Input rank | Income | Region | Population (mn) | GDP, PPP$ | GDP per capita, PPP$ | GII 2018 rank |
|---|---|---|---|---|---|---|---|
| **55** | **59** | **Upper middle** | **LCN** | **130.8** | **2,575.2** | **20,601.7** | **56** |

| | | Score/Value | Rank | |
|---|---|---|---|---|
| | **INSTITUTIONS** | **62.8** | **66** | |
| **1.1** | **Political environment** | **51.1** | **78** | |
| 1.1.1 | Political and operational stability* | 61.4 | 91 | |
| 1.1.2 | Government effectiveness* | 45.9 | 72 | |
| **1.2** | **Regulatory environment** | **59.0** | **84** | |
| 1.2.1 | Regulatory quality* | 47.2 | 61 | |
| 1.2.2 | Rule of law* | 31.4 | 97 | |
| 1.2.3 | Cost of redundancy dismissal, salary weeks | 22.0 | 94 | |
| **1.3** | **Business environment** | **78.4** | **37** | |
| 1.3.1 | Ease of starting a business* | 85.9 | 75 | |
| 1.3.2 | Ease of resolving insolvency* | 70.8 | 30 | ◆ |
| | **HUMAN CAPITAL & RESEARCH** | **33.4** | **54** | |
| **2.1** | **Education** | **43.5** | **76** | |
| 2.1.1 | Expenditure on education, % GDP | 5.2 | 38 | |
| 2.1.2 | Government funding/pupil, secondary, % GDP/cap | 15.6 | 79 | |
| 2.1.3 | School life expectancy, years | 14.3 | 66 | |
| 2.1.4 | PISA scales in reading, maths, & science | 415.7 | 55 | |
| 2.1.5 | Pupil-teacher ratio, secondary | 16.9 | 75 | |
| **2.2** | **Tertiary education** | **30.7** | **64** | |
| 2.2.1 | Tertiary enrolment, % gross | 38.2 | 72 | |
| 2.2.2 | Graduates in science & engineering, % | 25.5 | 27 | |
| 2.2.3 | Tertiary inbound mobility, % | 0.3 | 102 | ○ ◇ |
| **2.3** | **Research & development (R&D)** | **25.8** | **42** | |
| 2.3.1 | Researchers, FTE/mn pop. ⊖ | 244.2 | 74 | |
| 2.3.2 | Gross expenditure on R&D, % GDP ⊖ | 0.5 | 65 | |
| 2.3.3 | Global R&D companies, avg. exp. top 3, mn US$ | 49.0 | 29 | ◆ |
| 2.3.4 | QS university ranking, average score top 3* | 41.2 | 30 | ◆ |
| | **INFRASTRUCTURE** | **48.3** | **59** | |
| **3.1** | **Information & communication technologies (ICTs)** | **72.8** | **51** | |
| 3.1.1 | ICT access* | 54.9 | 79 | |
| 3.1.2 | ICT use* | 49.6 | 72 | |
| 3.1.3 | Government's online service* | 92.4 | 22 | ● ◆ |
| 3.1.4 | E-participation* | 94.4 | 17 | ● ◆ |
| **3.2** | **General infrastructure** | **31.9** | **76** | |
| 3.2.1 | Electricity output, GWh/mn pop. | 2,586.2 | 69 | |
| 3.2.2 | Logistics performance* | 46.2 | 50 | |
| 3.2.3 | Gross capital formation, % GDP | 22.5 | 70 | |
| **3.3** | **Ecological sustainability** | **40.1** | **54** | |
| 3.3.1 | GDP/unit of energy use | 11.6 | 34 | |
| 3.3.2 | Environmental performance* | 59.7 | 64 | |
| 3.3.3 | ISO 14001 environmental certificates/bn PPP$ GDP | 0.7 | 74 | |
| | **MARKET SOPHISTICATION** | **49.9** | **57** | |
| **4.1** | **Credit** | **37.3** | **62** | |
| 4.1.1 | Ease of getting credit* | 90.0 | 7 | ● ◆ |
| 4.1.2 | Domestic credit to private sector, % GDP | 35.5 | 87 | |
| 4.1.3 | Microfinance gross loans, % GDP ⊖ | 0.4 | 35 | |
| **4.2** | **Investment** | **32.8** | **110** | ○ |
| 4.2.1 | Ease of protecting minority investors* | 58.3 | 68 | |
| 4.2.2 | Market capitalization, % GDP | 34.4 | 44 | |
| 4.2.3 | Venture capital deals/bn PPP$ GDP | 0.0 | 69 | ○ |
| **4.3** | **Trade, competition, & market scale** | **79.5** | **8** | ● ◆ |
| 4.3.1 | Applied tariff rate, weighted avg., % | 1.2 | 12 | ● |
| 4.3.2 | Intensity of local competition† | 70.1 | 59 | |
| 4.3.3 | Domestic market scale, bn PPP$ | 2,575.2 | 11 | ● ◆ |

| | | Score/Value | Rank | |
|---|---|---|---|---|
| | **BUSINESS SOPHISTICATION** | **29.4** | **73** | |
| **5.1** | **Knowledge workers** | **35.7** | **68** | |
| 5.1.1 | Knowledge-intensive employment, % | 19.9 | 74 | |
| 5.1.2 | Firms offering formal training, % firms ⊖ | 50.8 | 20 | ● |
| 5.1.3 | GERD performed by business, % GDP ⊖ | 0.1 | 55 | |
| 5.1.4 | GERD financed by business, % | 20.7 | 66 | |
| 5.1.5 | Females employed w/advanced degrees, % | 8.8 | 74 | |
| **5.2** | **Innovation linkages** | **20.0** | **87** | |
| 5.2.1 | University/industry research collaboration† | 43.7 | 56 | |
| 5.2.2 | State of cluster development† | 53.8 | 39 | ◆ |
| 5.2.3 | GERD financed by abroad, % | 0.6 | 95 | ○ |
| 5.2.4 | JV–strategic alliance deals/bn PPP$ GDP | 0.0 | 81 | |
| 5.2.5 | Patent families 2+ offices/bn PPP$ GDP | 0.1 | 63 | |
| **5.3** | **Knowledge absorption** | **32.6** | **67** | |
| 5.3.1 | Intellectual property payments, % total trade | 0.1 | 104 | ○ |
| 5.3.2 | High-tech imports, % total trade | 17.0 | 10 | ● ◆ |
| 5.3.3 | ICT services imports, % total trade | 0.0 | 125 | ○ ◇ |
| 5.3.4 | FDI net inflows, % GDP | 3.1 | 54 | |
| 5.3.5 | Research talent, % in business enterprise ⊖ | 24.5 | 50 | |
| | **KNOWLEDGE & TECHNOLOGY OUTPUTS** | **25.5** | **50** | |
| **6.1** | **Knowledge creation** | **11.0** | **67** | |
| 6.1.1 | Patents by origin/bn PPP$ GDP | 0.5 | 76 | |
| 6.1.2 | PCT patents by origin/bn PPP$ GDP | 0.1 | 65 | |
| 6.1.3 | Utility models by origin/bn PPP$ GDP | 0.2 | 42 | |
| 6.1.4 | Scientific & technical articles/bn PPP$ GDP | 4.3 | 88 | |
| 6.1.5 | Citable documents H-index | 27.4 | 34 | ◆ |
| **6.2** | **Knowledge impact** | **36.7** | **65** | |
| 6.2.1 | Growth rate of PPP$ GDP/worker, % | 0.3 | 82 | |
| 6.2.2 | New businesses/th pop. 15–64 | 0.5 | 83 | ○ |
| 6.2.3 | Computer software spending, % GDP | 0.2 | 66 | |
| 6.2.4 | ISO 9001 quality certificates/bn PPP$ GDP | 2.9 | 77 | |
| 6.2.5 | High- & medium-high-tech manufactures, % | 0.5 | 11 | ● ◆ |
| **6.3** | **Knowledge diffusion** | **28.7** | **33** | ◆ |
| 6.3.1 | Intellectual property receipts, % total trade | 0.0 | 102 | ○ ◇ |
| 6.3.2 | High-tech net exports, % total trade | 15.0 | 9 | ● ◆ |
| 6.3.3 | ICT services exports, % total trade | 0.0 | 126 | ○ |
| 6.3.4 | FDI net outflows, % GDP | 0.7 | 61 | |
| | **CREATIVE OUTPUTS** | **29.2** | **55** | |
| **7.1** | **Intangible assets** | **41.4** | **62** | |
| 7.1.1 | Trademarks by origin/bn PPP$ GDP | 44.1 | 59 | |
| 7.1.2 | Industrial designs by origin/bn PPP$ GDP | 0.7 | 82 | |
| 7.1.3 | ICTs & business model creation† | 67.6 | 37 | ◆ |
| 7.1.4 | ICTs & organizational model creation† | 57.9 | 53 | |
| **7.2** | **Creative goods & services** | **32.1** | **22** | ● ◆ |
| 7.2.1 | Cultural & creative services exports, % total trade | 0.0 | 118 | ○ |
| 7.2.2 | National feature films/mn pop. 15–69 | 2.0 | 66 | |
| 7.2.3 | Entertainment & Media market/th pop. 15–69 | 7.5 | 40 | |
| 7.2.4 | Printing & other media, % manufacturing | 0.4 | 96 | ○ ◇ |
| 7.2.5 | Creative goods exports, % total trade | 9.6 | 1 | ● ◆ |
| **7.3** | **Online creativity** | **2.2** | **82** | |
| 7.3.1 | Generic top-level domains (TLDs)/th pop. 15–69 | 2.5 | 72 | |
| 7.3.2 | Country-code TLDs/th pop. 15–69 | 3.3 | 58 | |
| 7.3.3 | Wikipedia edits/mn pop. 15–69 | 3.4 | 93 | |
| 7.3.4 | Mobile app creation/bn PPP$ GDP | 0.7 | 66 | |

NOTES: ● indicates a strength; ○ a weakness; ◆ an income group strength; ◇ an income group weakness; * an index; † a survey question. ⊖ indicates that the economy's data are older than the base year; see Appendix II for details, including the year of the data, at http://globalinnovationindex.org. Square brackets [ ] indicate that the data minimum coverage (DMC) requirements were not met at the sub-pillar or pillar level.

Fuente: GII (2019).

Esta explicación del posicionamiento de México, en el *Global Innovation Index* **(GII, 2019)**, cita una combinación específica de factores o condiciones, algunos más complejos que otros (por ejemplo, capital humano e investigación o sofisticación de negocio). La explicación, como se observa, tiene el carácter de una trayectoria causal: siete factores que interactúan, para producir un resultado global y juntos explican el posicionamiento de México en el *Global Innovation Index* **(GII, 2019)**, los cuales esencialmente se tendrían que aumentar en los *factores sofisticación de negocio* (**29.4**, ítem 5) y *capital humano e investigación* (**25.5** ítem 6) del subíndice de innovación de entrada (*innovation input sub-index*).

Como casi todos los argumentos basados en el estudio de caso, el argumento de que esta combinación de *condiciones causales*, explican el posicionamiento de México en el *Global Innovation Index* **(GII, 2019)**, es un *argumento asimétrico*. Esto es, es una explicación positiva de una instancia, en este caso, el posicionamiento que guarda México en dicho reporte y no necesariamente pretende ser una explicación completa de todos los factores que la componen, ya sea para justificar dicho posicionamiento o incluso, su ausencia.

Por el contrario, si el argumento hubiera sido presentado de manera simétrica, la expectativa sería en orden de incrementar o decrementar dicho posicionamiento, satisfaciendo la trayectoria causal correspondiente. Sin embargo, pueden existir varias trayectorias causales para explicar el posicionamiento actual de México y evitar aquellas que sean menos favorables en sus indicadores o factores.

En el lenguaje del marco teórico de la investigación, la trayectoria causal que se exhibe en el reporte **GII (2019)** es parte de un conjunto más amplio de trayectorias causales dadas las diversas combinaciones de los factores que la componen, es decir, al ver todos los casos de posicionamiento como un conjunto, puede haber casos que muestren la misma trayectoria de México, pero puede haber muchos casos que muestren trayectorias causales alternativas. Este último hecho (muchas instancias que producen el resultado), no invalida la trayectoria causal del caso, como una combinación de ser

suficiente (pero no ser necesaria) de condiciones para el posicionamiento de México.

Tomando el caso como referencia, un investigador podría moverse en cualquiera de dos estrategias principales:

1. Encontrar otras instancias de posicionamiento y examinar hasta qué punto, están de acuerdo en mostrar la misma trayectoria causal del posicionamiento de México, con la misma combinación de los factores, es decir, hacer todo (o prácticamente todos) los casos posicionamiento que muestren los factores componentes. Esta estrategia, emplea la investigación cualitativa común, la estrategia de seleccionar la variable dependiente, un enfoque que es casi universal, pero erróneamente, condenado por los investigadores cuantitativos **(King et al., 1994)**.
2. La segunda estrategia, sería tratar de encontrar otras instancias de la trayectoria causal del posicionamiento de México en el reporte **GII (2019)** y examinar si estos casos también experimentaron con los *factores sofisticación de negocio* (ítem 5) y *capital humano e investigación* (ítem 6) del subíndice de innovación de entrada (innovation input sub-index). En esencia, el investigador seleccionaría los casos sobre la base de sus puntuaciones en la variable independiente. En este ejemplo, sin embargo, la variable independiente es una trayectoria causal con sus siete factores o condiciones principales, todas satisfechas. El objetivo de esta estrategia es, evaluar la trayectoria planteando la pregunta: ¿esta trayectoria causal, invariablemente (o al menos, con consistencia sustancial) conducirá a producir el posicionamiento que ha logrado México en el índice **GII (2019)**?

Ambas estrategias, tienen una naturaleza basada en el marco teórico y se ajustan a los dos enfoques generales de marco teórico, descritos previamente (Ver **Tabla 7.5**):

**Tabla 7.5. Estrategias de análisis en la búsqueda de comunalidades**

| Estrategia | Descripción |
|---|---|
| **1** | La primera estrategia es examinar casos que comparten un resultado dado (por ejemplo, CEOs de pymes innovadoras recién creadas) e intentar identificar sus *condiciones causales* compartidas (por ejemplo, la existencia de un modelo de negocio o el cálculo correcto de precio a su producto/servicio). El término *condición causal* se usa para referirse a un aspecto de un caso que sea relevante de alguna manera para la cuenta del investigador o explicación de algún resultado. Es decir, *identificar las* condiciones causales *compartidas por casos con el* mismo resultado, es muy apropiado para la evaluación de condiciones necesarias. |
| **2** | La segunda estrategia, consiste en examinar casos que comparten una *condición causal* específica o, más comúnmente, una condición específica como combinación de *condiciones causales*, y evaluar si estos casos exhiben el mismo resultado (por ejemplo, los países que combinan regulación sobre prácticas de, CEOs de pymes innovadoras y un bajo nivel de desarrollo económico ¿todos sufren de baja competitividad en sus resultados? En otras palabras, examinar casos con las mismas *condiciones causales* para ver si también, comparten el mismo resultado, es muy apropiado para la evaluación de condiciones suficientes, especialmente, combinaciones suficientes de las condiciones. Establecer las condiciones que son necesarias o suficientes es de gran interés para los investigadores que trabajan los aspectos macropolíticos y/o macrosociales. |

Fuente: Ragin (2008) con adaptación propia.

1. El primero, es un examen de si las instancias del resultado (posicionamiento de México en el **GII, 2019)** constituyen un subconjunto de instancias de una combinación de *condiciones causales* (es decir, la trayectoria causal del posicionamiento de México en el **GII, 2019).** Esta demostración, establecería que las *condiciones causales* en cuestión, son necesarias.
2. El segundo, es un examen de si las instancias de una combinación específica de *condiciones causales* (la trayectoria causal del posicionamiento de México en el **GII, 2019)** constituyen un subconjunto de instancias de un resultado (posicionamiento de México en el **GII, 2019**). Es decir, esta demostración sería establecer que la combinación de *condiciones causales*, es suficiente.

Por supuesto, ambas estrategias podrían usarse, y si ambos subconjuntos de relaciones fueran confirmados, entonces los dos conjuntos (el conjunto de casos con la trayectoria causal del posicionamiento de México y el conjunto de casos con el posicionamiento de México) coinciden. Si bien, puede parecer que las dos estrategias juntas constituyen un análisis correlacional, recuerde que las correlaciones son fuertes cuando hay varias instancias de caso nulo-nulo que carecen tanto de la trayectoria causal como del resultado. Ninguna de estas dos estrategias de investigación, depende de tales casos.

Uno de los aspectos más importantes del pensamiento *configuracional* es que se vincula directamente a casos, procesos causales y mecanismos causales **(Boswell y Brown, 1999)**. Es decir, generalmente existe una correspondencia directa en el trabajo *configuracional*, entre argumentos causales y análisis de niveles de caso. El argumento de que una combinación específica de *condiciones causales* genera algún resultado, dirige la atención no solo hacia casos específicos, como en el ejemplo del posicionamiento de México en el **GII (2019)**, sino también, hacia características específicas. Además, la combinación de las *condiciones causales* citadas por el investigador deben tener una coherencia interna, en otras palabras, la combinación debe tener sentido como una trayectoria causal. En última instancia, se puede observar la causalidad solo a nivel de caso; un argumento causal combinatorio proporciona orientación explícita respecto a qué observar en un caso empírico y, muy a menudo también, implica mecanismos causales específicos que vinculan, tanto a los diferentes factores reunidos con los indicadores de la naturaleza de sus conexiones que producen el resultado.

## Relación con el punto de vista cuantitativo

En la investigación cuantitativa convencional, las variables independientes son tratadas como causas analíticamente separables del resultado bajo investigación. Por lo general, se considera que cada variable causal tiene una autonomía o capacidad independiente para influir en el nivel, intensidad

o probabilidad de la variable dependiente. La mayoría de las aplicaciones de los métodos cuantitativos convencionales, suponen que los efectos de las variables independientes son tanto de tipo lineal como aditivo, lo que significa que el impacto de una determinada variable independiente sobre la variable dependiente, es la misma, independientemente de los valores de las otras variables independientes.

Las estimaciones de los *efectos netos*, suponen que el impacto de una determinada variable independiente, es la misma no solo a través de todos los valores de las otras variables independientes, sino también, en todas sus diferentes combinaciones **(Ragin, 2006a)**. Para estimar el efecto neto de una determinada variable independiente, el investigador compensa el impacto de las *condiciones causales* involucradas, restando de la estimación del efecto de cada variable, cualquier variación explicada en la variable dependiente, que comparte con otras variables causales. Este es el significado central de los *efectos netos*: el cálculo de la contribución no superpuesta o traslapada de cada variable independiente a la variación explicada en la variable dependiente. El grado de superposición o traslape, es una función directa de la correlación. Generalmente, cuanto mayor es la correlación de una variable independiente con respecto al resto, menor es su efecto neto.

Cuando se confrontan argumentos, que citan *condiciones causales* combinadas (por ejemplo, que se debe satisfacer una trayectoria causal de algún tipo), la recomendación habitual, es que los investigadores modelen combinaciones de condiciones como una serie de interacciones de efectos y prueben la importancia de la contribución incremental de la interacción estadística, a la variación explicada en la variable dependiente. Cuando hay interacción, el tamaño del efecto de un factor o variable independiente (por ejemplo, el subíndice de entrada de innovación) en una variable dependiente (índice de innovación global de un país) depende de los valores de uno o más factores o variables independientes (por ejemplo, el nivel de sofisticación de negocio). Siguiendo nuestro ejemplo del posicionamiento de México de acuerdo con el **GII (2109)**, un investigador podría argumentar se encuentra así, por la baja respuesta en el factor de conocimiento y tecnología producidos (*knowledge & technology outputs*).

Sin embargo, como se explica por **Ragin (1987)** y publicaciones relacionadas, las técnicas de estimación diseñadas para modelos lineales aditivos, a menudo se quedan cortos cuando la meta que se les asigna es la de estimar efectos complejos de interacción. Los requisitos de datos, por sí solos son sustanciales, especialmente cuando el objetivo es estimar interacciones de orden superior. Además, existen muchas controversias y dificultades alrededor del uso de cualquier variable que carezca de un punto cero significativo en los modelos de interacción multiplicativa **(Allison, 1977)**. En general, no es razonable esperar que las técnicas que están específicamente diseñadas para estimar los *efectos netos* de variables independientes en los modelos lineales aditivos realicen un buen trabajo al evaluar trayectorias causales, especialmente en situaciones donde existan, múltiples trayectorias causales en las que pueda estar implicado.

El desafío que plantea el pensamiento *configuracional,* es el d e a preciar las condiciones de causalidad no como factores adversarios en la lucha por explicar la variación en las variables dependientes, sino como posibles factores colaboradores en la producción de resultados. La cuestión clave, no es determinar qué variable es la más fuerte (es decir, la que tiene el mayor efecto neto) sino cómo se combinan las diferentes condiciones y si solo hay una combinación o varias combinaciones diferentes de condiciones (trayectorias *causales*) capaces de generar el mismo resultado. Una vez que se identifican estas combinaciones, es posible especificar los contextos que habilitan o deshabilitan las causas específicas. Por ejemplo, un investigador podría encontrar que la combinación de los factores de la **GII (2019)** de sofisticación de mercados (*market sophistication*) e infraestructura (infrastructure) con un alto nivel producen un incremento notable en el índice global de innovación (*Global Innovation Index*), solo en países emergentes. Por lo tanto, un contexto de impulso de mercado con tecnología, al combinarse, permiten un mejor posicionamiento de la innovación de un país emergente.

## Evaluación del grado de pertenencia en una configuración

Como alternativa a las pruebas de interacción estadística, considere abordar el problema de las *configuraciones* causales, como un problema de medición. Es decir, en lugar de probar los efectos de interacción contra los efectos aditivos en los modelos lineales. Como se recomienda en varios textos sobre métodos cuantitativos, simplemente, se mide el grado en el cual, están presentes las trayectorias causales específicas, a través de aplicar los conjuntos de datos, en casos relevantes. Con estos, el grado en el cual un caso exhibe una combinación de condiciones, está determinada por la condición que registra la más baja puntuación o el grado más bajo de expresión (es decir, la mínima puntuación de pertenencia). Este pensamiento, llamado del eslabón más débil sostiene, en efecto, que el grado en el cual, un caso expresa una combinación de condiciones (o configuración) es tan fuerte como lo es, su grado de expresión del elemento más débil. Los conjuntos de *datos difusos*, proporcionan una forma directa de operacionalizar este principio como una estrategia de medición **(Ragin, 2000)**, usando la intersección en los mismos.

Todo caso relevante, se calibra de la misma manera, con el grado de pertenencia en los conjuntos relevantes que van desde **0.0** (*no pertenencia plena*) a **1.0** (*pertenencia completa*). El grado de pertenencia en una combinación de miembros (es decir, *una configuración*) corresponde simplemente a la puntuación mínima (la más baja) de la pertenencia difusa entre las *condiciones causales* y otros aspectos que son combinados en una trayectoria causal. Por ejemplo, suponga que un investigador está interesado en conocer el grado, en el cual, diferentes países tienen la posibilidad de exhibir la misma trayectoria del país que está en el máximo puntaje del Reporte **GII (2019)**, en este caso Suiza (ver **Tabla 7.3**) para solo cuatro de siete factores que la componen. Ver **Tabla 7.6.**

**Tabla 7.6. Resumen de factores que componen al reporte GII (2019)**

| **Item** | **Siglas** | **Factor subíndice de innovación de entrada (*innovation input sub-index*)** |
|---|---|---|
| 1 | N | Instituciones (*Institutions*) |
| 2 | H | Capital humano e investigación (*Human capital & Research*) |
| 3 | I | Infraestructura (*Infrastructure*) |
| 4 | M | Sofisticación de mercado (*Market sophistication*) |
| 5 | B | Sofisticación de negocio (*Business sophistication*) |
| **Item** | **Siglas** | **Factor subíndice de innovación de salida (*innovation output sub-index*)** |
| 6 | K | Conocimiento y tecnología producidas (*Knowledge and technology outputs*) |
| 7 | C | Creatividad producida (*Creative outputs*) |

Fuente: GII (2019).

Descartando de inicio a instituciones **N** por no ser de interés. Una vez descartada y que estas condiciones sean calibradas como conjuntos de *datos difusos*, deberán mostrar el grado de pertenencia en cada uno de los casos y es posible usar estos para construir una medida simple del grado en que los casos combinan estos siete factores.

$$\mathbf{R_i = min\ (\ H_i\ ,\ I_i,\ M_i\ ,\ B_i\ )}$$

Donde **R**, es el grado de pertenencia en la *trayectoria causal* (la combinación de los cuatro factores): **min** indica la selección del más bajo valor de las cuatro puntuaciones difusas; ***i*** indica, que la fórmula se aplica a casos individuales.

Después de usar el conjunto de *datos difusos*, para calcular el grado en el cual, diferentes países se ajustan a esta trayectoria causal, el investigador podría evaluar la conexión entre la intensidad con la que los casos se exhiben (o se ajustan a) esta trayectoria causal y su grado de pertenencia en el resultado (es decir, el grado de pertenencia en el conjunto de países con factores similares al país más innovador: Suiza) también operacionalizado

como un conjunto difuso. Si la trayectoria causal de Suiza es una de varias trayectorias causales posibles (y por lo tanto, suficiente pero no necesario), entonces la expectativa es que los casos ajustados a esta trayectoria causal, deban constituir, un subconjunto de los casos como el de Suiza. Esta relación, si se trazara, aparecería como un triángulo superior izquierdo en el gráfico de grado de pertenencia en el conjunto de casos de países innovadores (**eje Y**) contra el grado de pertenencia a la trayectoria causal de Suiza dada (**eje X**).

## Comparando las trayectorias causales

El uso de conjuntos de *datos difusos*, para abordar trayectorias causales, es bastante flexible, y es una tarea sencilla formularla y comparar trayectorias causales de interés. También, es posible comparar versiones más simples de una trayectoria causal con versiones más complejas. Por ejemplo, un investigador puede argumentar que la necesidad de aislar el efecto del factor capital humano e investigación **H** ya que su marco teórico requiere probar lo que sucede si se retira para mejorar el posicionamiento de un país del índice global de innovación en el reporte **GII (2019)** y proponer una formulación alternativa que omita esta *condición causal* de la trayectoria, como caso especial. Así, se emite una versión simplificada, de tres condiciones de trayectoria, para el subíndice de innovación de entrada **GII (2019)** en este ejemplo:

$$\mathbf{R_i = min ( H_i , I_i, M_i , B_i )}$$

*Para comparar las dos trayectorias causales*:

1. El investigador medirá el grado de pertenencia a casos relevantes en ambas trayectorias (con cuatro y tres factores).
2. Evaluar la conexión del marco teórico de la investigación, entre las dos trayectorias por un lado y el resultado, por otro lado.

3. Para realizarlo, deberá usando las medidas de *consistencia* y *cobertura* del marco teórico.

En este ejemplo, la comparación de trayectorias, es entre una *condición causal* de cuatro vs. una de tres condiciones, donde esta última, es una versión simplificada y, por lo tanto, más inclusiva que la de cuatro condiciones. Tenga en cuenta, que el grado de pertenencia en la trayectoria causal de cuatro condiciones, es un subconjunto del grado de pertenencia del de tres condiciones. Es decir, el grado de pertenencia de cada caso en la trayectoria causal de cuatro condiciones debe ser menor o igual, que a su grado de pertenencia en la trayectoria de tres condiciones. Una forma simple, de entender esta propiedad matemática, es reconocer que con cuatro condiciones, hay una oportunidad adicional para que los casos reciban una baja puntuación en la trayectoria causal, debido a la inclusión del factor, capital humano e investigación **H.** Desde un punto de vista del marco teórico de la investigación, es cierto que los conjuntos componentes (como cuando se agrega la condición **H** a la combinación de **M, B, K, C**) produce subconjuntos a través de la intersecciones que se establezcan. La expectativa en este ejemplo es que la *trayectoria causal* de Suiza, es solo una de varias *trayectorias causales* para el logro del máximo posicionamiento en el reporte **GII (2019)**; por lo tanto, el grado de pertenencia en la trayectoria causal de Suiza, debe ser un subconjunto del grado de pertenencia del máximo posicionamiento en el reporte **GII (2019)**. El análisis del marco teórico de la investigación, por lo tanto, procedería como sigue:

1. Medir el grado de pertenencia, de cada caso en las cuatro condiciones de la trayectoria causal.
2. Medir el grado de pertenencia, de cada caso en las tres condiciones de la trayectoria causal.
3. Medir el grado de pertenencia de cada caso en el resultado, posicionamiento en el índice **GII (2019)**.
4. Evaluar la *consistencia* de la trayectoria causal de las cuatro condiciones como un subconjunto del resultado. Si es consistente, evalúe su *cobertura*.

5. Evaluar la *consistencia* de la trayectoria causal de las tres condiciones como un subconjunto del resultado. Si es consistente, evalúe su *cobertura*.
6. Compare los dos conjuntos de resultados.

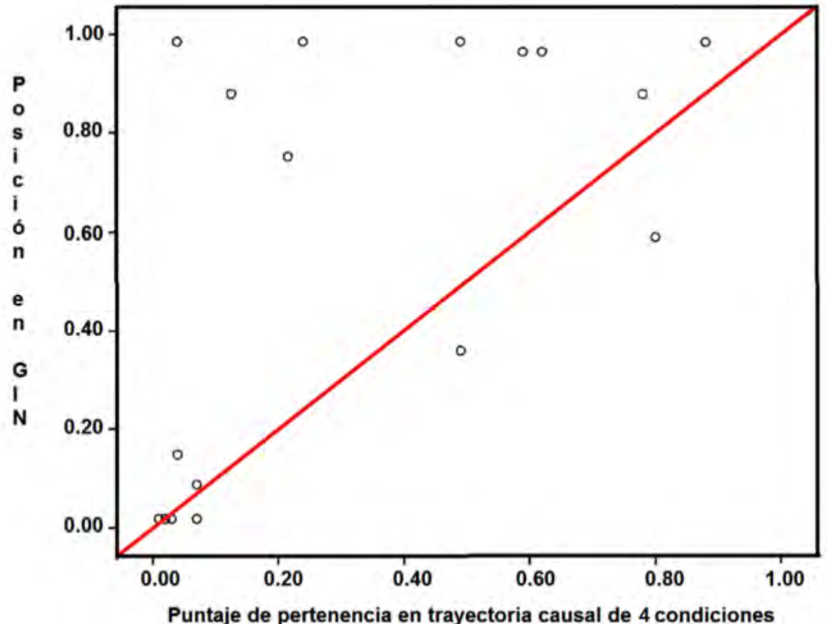
Como se ha señalado en capítulos previos, a menudo se produce una compensación entre la coherencia y *cobertura*. Esto es, debido a que los puntajes de pertenencia en las cuatro condiciones de la trayectoria causal, deben ser menores o iguales que los puntajes de pertenencia en la trayectoria causal de tres condiciones, la *consistencia* de la trayectoria causal de cuatro condiciones, en general será mayor o igual que la *consistencia* de la trayectoria causal de tres condiciones receta. Es posible construir una base de datos, en el cual, la *consistencia* del marco teórico de la trayectoria causal más simple, utilizando un subconjunto de componentes, sea más grande que la *consistencia* del marco teórico de una trayectoria más compleja, aunque esta situación, es muy rara empíricamente.

Lo contrario sucede con la *cobertura*. Debido a que las puntuaciones de pertenencia en la trayectoria causal de tres condiciones, deben ser mayores o iguales a puntuaciones de pertenencia en la trayectoria causal de cuatro condiciones, posteriormente la *cobertura* de la trayectoria causal de tres condiciones, generalmente, será mayor o igual a la *cobertura* de la trayectoria causal de siete condiciones. Los gráficos, se ilustran en las **Gráficas 7.1 y 7.2**, como casos hipotéticos. El puntaje de *consistencia* también permite la interpretación del puntaje de *cobertura*, que es **0.53**. La **Gráfica 7.1**, muestra la trayectoria causal de cuatro condiciones vs. el resultado hipotético (posicionamiento en el reporte GNI, 2019), junto con la *consistencia* y las medidas de *cobertura*.

La **Gráfica 7.2** proporciona la misma información usando trayectoria causal de tres condiciones. Observe que en la primera gráfica, la mayoría de los puntos están consistentemente por encima de la diagonal; los pocos que se alejan no están muy por debajo la diagonal, por lo tanto, este gráfico registra una puntuación de alta *consistencia*, **0.93**, indicando que la evidencia apoya la afirmación de que la pertenencia en la trayectoria causal

de cuatro condiciones, es un subconjunto de la pertenencia en el resultado, lo que indica a su vez, que sería razonable intentar una interpretación del ser suficiente causal para esta trayectoria causal. Esta, relativamente alta *consistencia* también permite la interpretación del puntaje de *cobertura,* el cual es **0.53**.

**Gráfica 7.1. Puntuación de pertenencia de posicionamiento en GII (2019) vs. la trayectoria causal de 4 condiciones**




Notas: Consistencia marco teórico: 0.93 ; Cobertura marco teórico: 0.53.

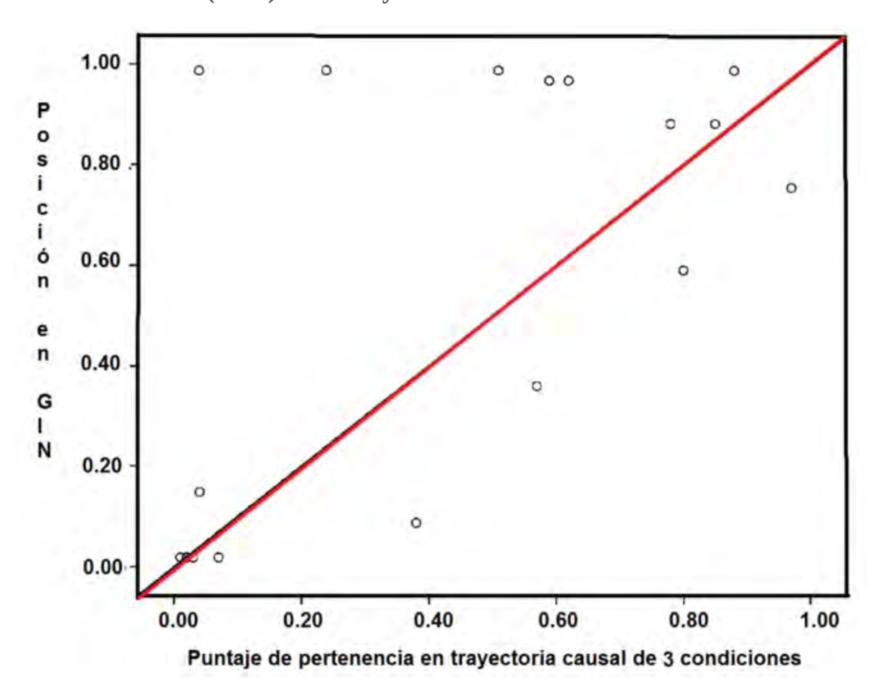
Fuente: Ragin (2008) con adaptación propia.

Por lo tanto, la *cobertura* del resultado para las cuatro condiciones, la trayectoria causal es sustancial y representa más de la mitad de la suma, de las pertenencias en el resultado. Por el contrario, la trayectoria causal de tres condiciones, tiene más casos debajo de la diagonal y por lo tanto, menor puntuación de *consistencia,* **0.79.** La brecha entre los dos puntajes **(0.93 a 0.79)** indica un aumento sustancial en la *inconsistencia* cuando la condición

*Instituciones* **I**, se elimina de la trayectoria causal. En consecuencia, este puntaje de *consistencia* límite **(0.79)** indica que no sería pertinente inter-pretar la medida de*cobertura* informada para la **Grá ica 7.2 (0.66).**

El procedimiento recién descrito, se puede extender para que los investigadores puedan evaluar todas las versiones posibles de una trayectoria causal dada (es decir, todos los subconjuntos posibles de los componentes especificados en la versión completa de la trayectoria causal). En el ejemplo que se muestra en la **Grá ica 7.2**, solo se han eliminado las *condición causal*: *instituciones* **N** (se declara por parte del investigador de nulo interés) y se aísla *capital humano e investigación* **H**, produciendo una sola trayectoria causal de tres condiciones para el subíndice de innovación de entrada **GII (2019).**

**Gráfica 7.2. Grá ica de puntuación de pertenencia de posicionamiento en GII (2019) vs. la trayectoria causal de 3 condiciones**




Notas: Consistencia marco teórico: 0.79 ; Cobertura marco teórico: 0.66.

Fuente: Ragin (2008) con adaptación propia.

Además de la trayectoria causal de cuatro condiciones, el investigador considera bajo su marco de experiencia sustantiva, que otras más como: cuatro de tres condiciones de trayectoria causal (cuando se usa la trayectoria causal de cuatro condiciones como inicio), seis trayectorias causales de dos condiciones y cuatro trayectorias causales de una condición. Utilizando métodos difusos, el grado de pertenencia de casos en cada una de estas quince trayectorias causales posibles, pueden ser evaluadas de cada una, su *consistencia* y *cobertura* como un subconjunto del resultado. De esta manera, los investigadores pueden comparar, versiones más simples de una trayectoria causal con versiones más complejas. Si una versión más simple, tiene una *consistencia* comparable pero de mayor *cobertura* que el de versión más compleja, entonces es preferible.

Como un ejemplo de este tipo de análisis, suponga que el mismo investigador del caso anterior, desea examinar las quince trayectorias causales, las cuales han revelado que los grados de pertenencia en una trayectoria causal de dos condiciones, digamos la combinación de factores como *la sofisticación de mercado* **M** y *negocio* **N,** forman un subconjunto consistente al resultado, posicionamiento del país al índice **GII (2019)**, con un puntaje de *consistencia* de **0.90** y una *cobertura* de **0.71**. El investigador, bien pudiera preferir este argumento como explicación más *parsimoniosa*, que el argumento original, la trayectoria causal de cuatro condiciones. Mientras que una pequeña puntuación de *consistencia* del marco teórico, ha sido sacrificada (**0.93** para la trayectoria causal de cuatro condiciones vs. **0.90** para la trayectoria causal de dos condiciones), se manifiesta un aumento sustancial en la *cobertura* del marco teórico de la investigación (**0.71** para la trayectoria causal de dos condiciones vs. **0.53** para la trayectoria causal de siete condiciones). Una puntuación de *cobertura* de **0.71**, indica que la trayectoria causal de dos condiciones, representa el **71%** de la suma de puntuaciones de pertenencia en el resultado. Aún así, es importante evitar un uso abusivo de la *parsimonia* ya que hay tendencia a discusiones con comparativos de índices de mínima diferencia, siendo trivial su exposición, por lo que se sugiere siempre hacer una tabla que facilite la comparación y a partir de ahí, hacer su discusión. Ver **Tabla 7.7.**

**Tabla 7.7. Trayectorias causales a casos hipotéticos GII (2019)**

| Casos de trayectoria causal | Marco teórico | |
|---|---|---|
| | Consistencia | Cobertura |
| Cuatro condiciones | 0.93 | 0.53 |
| Tres condiciones | 0.79 | 0.66 |
| Dos condiciones | 0.90 | 0.71 |

Fuente: Ragin (2008) con adaptación propia.

El impulso científico, tendería a hacer una discusión *parsimoniosa* a partir de la trayectoria causal de tres condiciones. Pero considere el hecho, que desde una perspectiva *configuracional*, el trayectoria causal de cuatro condiciones podría tener más sentido. Es decir, al usarlo como guía para comprender los casos, la trayectoria causal de cuatro condiciones podría ofrecer una visión más completa, conectarse mejor a los procesos causales observados, y ofrecer una mejor base para comprender los mecanismos causales de la investigación. En resumen, si una explicación más compleja, proporciona puntuaciones similares a una explicación más *parsimoniosa*, la explicación más compleja podría preferirse a la explicación más *parsimoniosa* por motivos sustantivos y teóricos.

Dado que la *consistencia* es una proporción, estas pruebas se pueden aplicar fácilmente. Por ejemplo, como se muestra en **Ragin (2000)**, es posible establecer puntuaciones de *consistencia* de referencia y un nivel alfa, antes de hacer cálculos del resto de las *consistencias*, y posteriormente descubrir, qué *consistencias* observadas son significativamente mayores al valor de referencia, dado el número de casos incluidos en el estudio.

Por ejemplo, puede preguntarse ¿es una puntuación de *consistencia* de **0.93** significativamente mayor que una puntuación de referencia de **0.75**, usando un nivel alfa de **0.05** y un **N** de **50**? es sencillo de calcular. Es también posible determinar, si un puntaje de *consistencia* observado es significativamente mayor que otra puntuación de *consistencia* observada, para un determinado número de casos y un alfa dado. Esta misma prueba, se puede

calcular para dos puntuaciones de *cobertura* observados, que también son proporciones. Esta unión del análisis del marco teórico de la investigación y el análisis probabilístico es útil, siendo alentadas también, otras posibilidades a explorar. El énfasis aquí, sin embargo, se trata de una comprensión descriptiva de la *consistencia y cobertura*. Finalmente, es importante señalar nuevamente, que la prueba real de cualquier *trayectoria causal* es qué tan bien se acopla con el conocimiento del caso. Una parte importante de cualquier análisis del marco teórico es su aplicación al estudio de los casos. Como se señaló en el principio, el pensamiento *configuracional* en general y la idea de las trayectorias causales específicamente, es facilitar el diálogo entre los análisis de cruce de casos y el análisis de caso. Cuando son formulados como trayectorias causales (no como *efectos netos* de variables independientes) la relación entre los hallazgos analíticos con los casos empíricos, es altamente robusto.

## Resumen fsQCA para cómo evaluar una trayectoria causal

Se listan las etapas de acuerdo con la **Tabla 7.8.**

**Tabla 7.8. Resumen fsQCA para cómo evaluar una trayectoria causal**

| Item | Etapa | Descripción |
|---|---|---|
| 1 | **Definición de la trayectoria causal** | Por lo general:<br>**a.** La trayectoria causal es una combinación específica de condiciones,<br>**b.** El primer objetivo es evaluar la conformidad de cada caso a esta combinación específica. Por ejemplo, un investigador podría especificar una trayectoria causal que combine con cuatro condiciones (factores). |
| 2 | **Determinación grado de pertenencia por caso** | Parla lograrlo:<br>**a.** Determine el grado de pertenencia de cada caso en cada una de las condiciones incluidas en la trayectoria causal, concibiendo cada condición como conjunto de *datos difusos* (*fuzzy-sets*).<br>**b.** Si es necesario, utilice los procedimientos detallados a lo largo del libro, para transformar variables de escala de intervalo y conjunto de *datos difusos* (*fuzzy-sets*) bien calibrados. |

| | | |
|---|---|---|
| 3 | **Cálculo de grado de pertenencia por trayectoria causal** | Para lograrlo:<br>**a.** Calcule el grado de pertenencia de cada caso de *trayectoria causal*.<br>**b.** Para hacer esto, primero recupere su conjunto de datos (con conjuntos difusos bien calibrados) en **fsQCA**.<br>**c.** Con su conjunto de datos en la ventana de hoja de cálculo de datos, haga clic en Variables (*Variables*).<br>**d.** Posteriormente haga clic en Calcular (*Compute*) y aparecerá el cuadro de diálogo Calcular Variable (*Compute Variable*).<br>**e.** Nombre el objetivo difuso (*Name Target Fuzzy-Set*), seleccionando un nombre simple (dos a ocho caracteres), utilizando caracteres alfanuméricos estándar y sin espacios, guiones o signos de puntuación (por ejemplo, *trayectoria1*). |
| 4 | **Habilitando las funciones** | Para logralo:<br>**a.** Haga clic en *fuzzyand (x,...)* En el menú Funciones (*Functions*) y posteriormente haga clic en el flecha hacia arriba que está al lado de la palabra Funciones (*Functions*).<br>**b.** A continuación, aparecerá *fuzzyand ()* en el campo Expresión (*Expression*) del cuadro de diálogo. La función *fuzzyand* realiza una *intersección difusa*, lo que significa que selecciona el más bajo puntaje de pertenencia para cada caso de los conjuntos de *datos difusos* (*fuzzy-sets*) que se enumeran dentro los paréntesis. Con la *intersección difusa*, el grado en que un caso se muestra que una *trayectoria causal* dada, es tan fuerte como la expresión de su factor constituyente más débil. |
| 5 | **Editar la intersección difusa** | Para lograrlo:<br>**a.** Edite la expresión de intersección para que fuzzyand () se convierta en algo así como *fuzzyand (fvar1, fvar2, fvar3, fvar4),* donde *fvar1* a *fvar4* son los nombres de las variables de las condiciones que componen la *trayectoria causal* (en este ejemplo, hay siete condiciones), utilizando nombres de variables de conjunto difuso en el existente archivo de datos.<br>**b.** Haga clic en Aceptar (*OK*). |

| | | |
|---|---|---|
| 6 | **Comprobación en la hoja de cálculo** | Para lograrlo:<br>**a.** Compruebe la hoja de cálculo de datos, para asegurarse de que salió como esperaba.<br>**b.** Es posible clasificar el grado de pertenencia en el nuevo cálculo de la trayectoria causal, en orden descendente o ascendente utilizando el menú desplegable.<br>**c.** Haga clic en cualquier caso en la columna que desea ordenar, posteriormente haga clic en Casos (*Cases*), y posteriormente Ordenar ascendente o descendentemente (*Sort Ascending o Descending*). |
| 7 | **Evaluación** | Para lograrlo:<br>**a.** Una forma sencilla de evaluar la conexión entre el marco teórico y la *trayectoria causal* recién calculada y un resultado difuso, es usar el procedimiento de graficación XY.<br>**b.** Desde la ventana de la hoja de cálculo de datos, haga clic en Gráficos (*Graphs*), posteriormente Difuso (*Fuzzy*), posteriormente Diagrama X Y (*XY Plot*).<br>**c.** Especifique los conjuntos difusos que se trazarán en los ejes X y Y haciendo clic en las flechas hacia abajo adyacentes y posteriormente haciendo clic en los nombres de variables relevantes.<br>**d.** Para el eje Y, haga clic en el conjunto de *datos difusos* (*fuzzy-sets*) para el grado de pertenencia en el resultado.<br>**e.** Para el eje X, haga clic en el grado de pertenencia en la *trayectoria causal*.<br>**f.** Especifique una variable ID (*Case Id Variable*) de caso opcional para que el caso o los casos, que residen en puntos específicos de la gráfica, pueden ser identificados. |
| 8 | **Graficar** | Para lograrlo:<br>**a.** Haga clic en Trazar (*Plot*).<br>**b.** Si la expectativa es que la pertenencia en la *trayectoria causal* es suficiente pero no necesario para ser miembro del resultado, entonces las puntuaciones deben estar sobre o encima de la diagonal principal de la gráfica.<br>**c.** Los números que se muestran en el cuadro, que están encima de la esquina superior izquierda de la gráfica muestran el grado de *consistencia* de la gráfica con la relación de subconjunto ( $X_i <= Y_i$ ).<br>**d.** El número que está, en el cuadro debajo de la esquina inferior derecha esquina de la gráfica, es el grado de *cobertura* del marco teórico del resultado por la *trayectoria causal* especificada . |

Fuente: Ragin (2000) con adaptación propia.

CAPÍTULO 8.

# Efectos netos en fsQCA

Los métodos de análisis de datos convencionales, como la regresión múltiple, conforman la columna vertebral, de la mayoría de las técnicas de la investigación cuantitativa en las ciencias sociales, con orientación a la administración de la innovación de hoy. No debería sorprender que lo hagan, ya que son considerados por muchos, por ser de los más rigurosos, más disciplinados y más científicos de los métodos analíticos disponibles para los investigadores. Si los resultados de la investigación social tienen, un impacto en la sociedad en general, tales hallazgos deben producirse utilizando los métodos analíticos más rigurosos posibles.

Si bien, los métodos cuantitativos convencionales son claramente rigurosos, están organizados alrededor de un tipo específico de rigor. Es decir, tienen su propio rigor y su propia disciplina, no un rigor universal. Existen varias características de los métodos cuantitativos convencionales, que los hacen rigurosos y por lo tanto, valiosos como herramientas analíticas. Así, este capítulo, se enfoca en un solo aspecto clave: el hecho de que normalmente se centran en la tarea de estimar los *efectos netos* de las variables independientes en los resultados. Este aspecto central, ya al caracterizarlos desde el punto de vista de los *efectos netos*, porque justamente esto es lo que limita a los métodos convencionales, su utilidad. Si bien, es importante evaluar la importancia relativa de las variables independientes, esta tarea no debe ser el foco exclusivo de la analítica de las ciencias sociales. En cambio, además de evaluar los *efectos netos*, los investigadores deben examinar cómo se combinan las diferentes *condiciones causales* para producir un resultado dado. Este capítulo presenta una crítica de los *efectos netos*, enfocándose en sus limitaciones, y describe los contrastes clave entre el pensamiento de *efectos netos* y el enfoque *configuracional* del análisis cualitativo comparativo (**QCA**). El capítulo 9 proporciona un ejemplo, utilizando un gran conjunto de datos **N**. Es importante señalar, que el argumento presen-

tado aquí, no es que las técnicas analíticas convencionales sean defectuosas o presenten fallas; de hecho, son rigurosas y muy potentes. Más bien, el argumento es que no están bien diseñadas y adecuadas para analizar la *complejidad causal*. De hecho, la evaluación de *efectos netos*, requieren que el investigador asuma que la causalidad no es complicada.

## La corriente de efectos netos

En lo que se ha convertido la práctica de la ciencia social normal, es en lo que los investigadores ven como su principal tarea, la de evaluar la importancia relativa de las variables causales extraídas de teorías diversas. En una situación ideal, las teorías relevantes enfatizan diferentes variables y aclaran, sin ambigüedades, declaraciones acerca de cómo estas variables están conectadas a resultados empíricos relevantes. En la práctica, sin embargo, la mayoría de las teorías en las ciencias sociales orientadas a la administración de la innovación, son vagas cuando se trata de especificar ambas, las *condiciones causales* y los resultados, tendiendo a ser aún más reservados cuando se trata de indicar cómo las *condiciones causales* están conectadas a los resultados (por ejemplo, especificando las condiciones que deben cumplirse, para que la variable causal dada tenga su impacto).

Por lo general, los investigadores solo pueden desarrollar listas de *condiciones causales*, potencialmente relevantes, basadas en el amplio panorama de los fenómenos sociales que encuentran en teorías diversas. La clave de la tarea analítica, generalmente se ve como la de evaluar la importancia relativa de las variables relevantes.Si las variables asociadas con una teoría particular demuestran ser los mejores predictores del resultado (es decir, los mejores explicadores de su variación), entonces esta teoría gana la competencia o interacción factorial. Esta forma de realizar análisis cuantitativos, es el procedimiento predeterminado en las ciencias sociales orientadas a la administración de la innovación hoy en día, una en que los investigadores retroceden de tiempo en tiempo, a menudo, por falta de conocimiento de una alternativa clara.

En el enfoque de los *efectos netos*, las estimaciones de los efectos de las variables independientes, se basan en el supuesto de que cada variable, por sí misma, es capaz de influir en el nivel o la probabilidad del resultado. Mientras es común tratar causal e independiente como sinónimos modificadores de la palabra variable, el significado central de independiente se relaciona con la noción de capacidad autónoma. Específicamente, cada variable independiente se supone que es capaz de influir en el nivel o la probabilidad de resultado, independientemente de los valores o niveles de otras variables (es decir, independientemente de los contextos variados definidos por estas variables). Las estimaciones de los *efectos netos* asumen así la aditividad, que el impacto neto de una determinada variable independiente en un resultado es la misma en todos los valores de las otras variables independientes y sus diferentes combinaciones.

Al estimar el *efecto neto* de una variable dada, el investigador compensa el impacto de *condiciones causales* competidoras, restando de la estimación del efecto de cada variable, cualquier variación explicada en la variable dependiente, que comparte con otras variables causales. Este es el significado central de los *efectos netos*: el cálculo de la contribución no superpuesta o de punto de cruce, de cada variable independiente a la variación explicada en el resultado. El grado de superposición es una función directa de la correlación. Generalmente, a mayor la correlación de una variable independiente con las restantes contribuyentes, menor es su efecto neto. Existe una afinidad subyacente importante, entre teoría pobremente especificada y la corriente de los *efectos netos*. Cuando las teorías son débiles, ofrecen solo caracterizaciones generales de fenómenos sociales y no abordan la *complejidad causal*. Cuando hay claras especificaciones de contextos relevantes y las condiciones de alcance son raras, como la consideración de, cómo las *condiciones causales* pueden modificar entre ellas, la relevancia o el impacto (es decir, cómo pueden mostrar no aditividad).

Los investigadores serán afortunados, si pueden derivar listas coherentes de *condiciones causales* potencialmente relevantes, de la mayoría de las teorías en las ciencias sociales, porque la teoría típica ofrece muy poca orientación específica. Este vacío, se llena con modelos lineales y

aditivos, con énfasis en estimar *efectos netos* genéricos. Los investigadores, a menudo declaran que estiman modelos aditivos lineales porque son los modelos más simples posibles y hacen los menos supuestos sobre la naturaleza de la causalidad.

Desde este punto de vista, la aditividad (y por lo tanto, la simplicidad causal) es el estado predeterminado; cualquier análisis de no aditividad requiere autorización teórica explícita, la cual, generalmente falta. Este énfasis común, en el cálculo de los *efectos netos* se ajusta con la idea de que el objetivo principal de la investigación social es evaluar el poder explicativo, relativo de las variables asociadas a las teorías en discusión. Los análisis de *efectos netos*, proporcionan evaluaciones cuantitativas explícitas de la variación explicada sin solapamiento, que se puede acreditar a cada variable de la teoría. Por lo tanto, el cálculo de los *efectos netos* aporta una fuerte base para la adjudicación de la teoría, proporcionando mayor justificación para el uso de estos métodos.

A menudo, sin embargo, las teorías no se contradicen, unas a otras, directamente y por lo tanto no compiten realmente. Después de todo, la típica teoría de ciencia social es aún pequeña e incipiente en el amplio panorama de fenómenos que se presentan. El uso de la enfoque de *efectos netos*, por lo tanto, puede crear la apariencia de la adjudicación de la teoría en investigaciones, en las que tal adjudicación puede no ser necesaria o incluso posible.

## Problemas con la corriente de efectos netos

Varios problemas están asociados con el enfoque de *efectos netos*, especialmente cuando se trata como el medio exclusivo, o incluso, el principal de generación de conocimiento científico social útil. Estos incluyen ambos, problemas prácticos y conceptuales. Un problema práctico fundamental, es el hecho de que la evaluación de los *efectos netos* dependen de la especificación del modelo. La estimación de un efecto neto de la variable independiente, se ve influido por sus correlaciones con el resto de las variables contribuyentes. Limíte el número de variables contribuyentes correlacio-

nadas, y una variable elegida puede tener un efecto neto sustancial en el resultado; apílelos, y su efecto neto puede reducirse a cero.

La dependencia de la especificación de la estimación de los *efectos netos*, es bien conocido, lo que explica por qué los investigadores cuantitativos están completamente orientados en la importancia de la especificación correcta. Sin embargo, una especificación correcta, es dependiende de un marco teórico de la investigación y de un fondo sustancial conocimiento, que generalmente carecen de aplicaciones de métodos de *efectos netos*.

La importancia de la especificación del modelo, es evidente en muchos análisis del conjunto de datos, como el usado en el caso hipotético del capítulo 11, en el que se reporta un muy fuerte efecto neto de las puntuaciones de prueba de factores que componen las capacidades de innovación de empresas clúster sobre los resultados de, evitar disminución de capacidades creativas. Se encuentra que cuanto mayor sea el puntaje de capacidades de innovación menor será la estimación del efecto neto de evitar la disminución de capacidades creativas. Por el contrario, si se utilizan los mismos datos en una regresión logística, se encontrará un débil efecto de las capacidades de innovación de empresas clúster sobre evitar la disminución de capacidades creativas. La diferencia entre ambos análisis, se puede explicar en las bases de diseño, por ejemplo, el primer análisis solamente permite que algunas variables interactúen, tal vez una o dos, mientras que el segundo análisis (regresión logística), son todas.

Por lo tanto, cabe preguntarse ¿cuál medición de la capacidad de innovación de empresas clúster es la correcta? La respuesta depende de qué especificación se considera correcta, por lo tanto, los debates sobre los *efectos netos*, a menudo se estancan en desacuerdos sobre la especificación del modelo. Mientras que los científicos sociales orientados a la administración de la innovación, tiendan a pensar que tener una gran cantidad variables es mejor que tener pocas, como en el análisis de regresión logística mencionada, el tener demasiadas variables independientes, puede producir errores de especificación paralizantes y producir resultados no interpretables **(Achen, 2005)**.

Un problema práctico relacionado, es el hecho de que muchas de las variables independientes que interesan a los científicos sociales orientados a la administración de la innovación, están altamente correlacionadas con entre sí y, por lo tanto, solo pueden tener modestos efectos no superpuestos en un resultado dado. Una vez más, la controversia del caso de capacidades de innovación es un buen ejemplo. Los puntajes de ingresos por innovaciones con educación de posgrado del CEO, se encuentran fuertemente correlacionados, al igual que estas dos variables con una variedad de otras potencialmente *condiciones causales relevantes* (tecnología,empresa altamente innovadora, tecnología media, CEO educación de posgrado, etcétera). Debido a que las desigualdades de los factores que componen las capacidades de innovación se superponen, los puntajes de los casos en las variables independientes las tienden a agruparse: altas puntuaciones de las capacidades de innovación tienden a sesgarse al ser los mejores en: tecnología, empresa altamente innovadora, tecnología media, CEO educación de posgrado, etcétera.

Por supuesto, estas correlaciones están lejos de ser perfectas; por lo tanto, es posible exprimir estimaciones de los *efectos netos* de estas variables independientes, fuera de los datos. Aún así, el abrumador patrón empírico es una de las causas confusas, de condiciones de agrupaciones (clúster) de favorables vs. desfavorables, no de las analíticamente separables, variables independientes. Una cuestión que los científicos sociales con orientación a la administración de la innovación saben sobre las desigualdades, es que al superponerse, se refuerzan. Es su naturaleza superpuesta, lo que les da su fuerza y durabilidad. Dado este rasgo característico de los fenómenos sociales, para los científicos sociales cuantitativos que se basan casi exclusivamente en técnicas que privilegian la estimación del efecto neto único y por separado de cada variable causal, parece algo muy contradictorio.

De manera más general, si bien es útil examinar las correlaciones entre las variables independientes, para ajustar la estimación de los *efectos netos* en consecuencia, también es útil estudiar los casos de manera integral, como las *configuraciones* específicas de los atributos. En esta vista, los

casos combinan diferentes características causalmente relevantes de diferentes maneras, y es importante para evaluar las consecuencias de estas diferentes combinaciones. Considere, por ejemplo, lo que se necesita para evitar la disminución de creatividad en una empresa clúster a través de la pregunta: ¿hace la educación universitaria de posgrado del CEO de una empresa clúster, con tecnología alta, que tenga altos ingresos por innovación? Probablemente no, o al menos no con gran diferencia. Pero la educación universitaria de posgrado del CEO puede ser la gran diferencia para una empresa del clúster con tecnología media para mejorar sus ingresos por innovación. Al examinar los casos como *configuraciones*, es posible realizar evaluaciones específicas del contexto, que están delimitados circunstancialmente. Las evaluaciones de este tipo implican preguntas sobre las condiciones que habilitan o deshabilitan conexiones específicas entre causas y resultados **(Amenta y Poulsen, 1996)**.

Así, cabe preguntarse: ¿bajo qué condiciones importan los puntajes de las pruebas cuando se trata de evitar disminución de capacidad creativa?, ¿bajo qué condiciones importa el posicionamiento de mercado?, ¿son estas conexiones diferentes por el género del CEO? Este tipo de preguntas están fuera del alcance de los análisis convencionales de *efectos netos*, por aquellos enfoques que se centran en la tarea de estimar el contexto independiente de los *efectos netos*.

Dichas evaluaciones de configuración, son directamente *relevantes* para los debates como de políticas de innovación en general. El discurso político, a menudo se centra en categorías y tipos de personas (o casos), no sobre las variables y sus *efectos netos* a través de poblaciones heterogéneas. Considere, por ejemplo, frases como empresas verdaderamente desfavorecidas, innovaciones de pobre impacto social o innovaciones como fuente de bienestar. En general, tales categorías abarcan combinaciones de características. Considere también el hecho, de que la política de innovación, tiende a ser en los últimos años y fundamentalmente de intervención social. Si bien podría ser bueno saber que la educación del CEO, en general, disminuye las probabilidades de disminuir las capacidades de innovación de la empresa,

desde una perspectiva política de innovación es mucho más útil saber cuál es el impacto social que conllevan las innovaciones que produce y bajo qué condiciones la educación del CEO con sentido social tiene un impacto decisivo, para beneficiar a su comunidad por las innovaciones introducidas.

Típicamente, los *efectos netos* se calculan a través de muestras extraídas de las poblaciones. No se basan en comparaciones estructuradas y enfocadas **(George, 1979, pp. 54–55)** utilizando tipos y categorías de casos específicos. Finalmente, mientras que el cálculo de los *efectos netos*, ofrece evaluaciones sucintas del poder explicativo relativo de las variables extraídas de diferentes teorías, la adjudicación entre el contenido del marco teórico, no es una preocupación central de la política de innovación con carácter social. ¿Qué teoría prevalece en el marco teórico de la investigación? Para explicar la variación en los resultados como evitar la disminución de capacidad creativa es principalmente un pregunta académica. El tema que es importante para la sociedad en general, especialmente cuando el objetivo es la intervención, es determinar qué *condiciones causales*, son decisivas y en qué contextos, independientemente de la teoría de la que se extraen las condiciones.

Para resumir, el enfoque de *efectos netos*, aunque potente y riguroso, está limitado. Está restringido por su propio rigor, porque su fuerza es también su debilidad. Es particularmente desventajoso cuando se usa para estudiar combinaciones de características del caso, especialmente desigualdades superpuestas. Dado estos inconvenientes, es razonable explorar un enfoque alternativo, uno con fortalezas que difieren de las de los métodos de *efectos netos*. Específicamente, el enfoque de *efectos netos*, con su fuerte énfasis en calcular el efecto no contaminado de cada variable independiente para aislar su impacto independiente, puede ser contrarrestado y complementado con un enfoque que considera, explícitamente, combinaciones y *configuraciones* de aspectos del caso.

## Cambiando el enfoque a tipos de casos

Subyacente a que la amplia extensión de la metodología científica social, orientada a la administración de la innovación, se encuentra en el continuum que se extiende, desde pequeñas muestras **N**, tratadas con técnicas cualitativas hasta las grandes muestras **N**, tratadas con técnicas cuantitativas. En general, los científicos sociales orientados a la innovación, deploran el amplio abismo que separa los dos extremos de este continuum, pero generalmente se adhieren a uno solo de estos extremos cuando inician una investigacion.

Con **QCA**, sin embargo, es posible traer algo del espíritu y la lógica de la investigación orientada a casos para la investigación de grandes muestras **N**. Esta técnica, ofrece a los investigadores herramientas para estudiar casos como *configuraciones* para explorar las conexiones entre combinaciones de condiciones de causalidad relevantes y resultados. Estudiando las combinaciones de *condiciones causales*, es posible desentrañar las condiciones y contextos que habilita o deshabilitan conexiones específicas (por ejemplo, entre preparación del CEO y el evitar pérdida de creatividad). El punto de partida de **QCA** es el principio de que los casos deben ser vistos, en términos de las combinaciones de condiciones de causalidad relevantes que estos exhiban. Para representar combinaciones de *condiciones causales*, los investigadores usan *tablas de verdad*, que enumeran las combinaciones lógicamente posibles de condiciones de causalidad especificadas por el investigador, ordenando los casos según muestren las combinaciones y enumeran un valor de resultado (normalmente codificado en, verdadero o falso) para cada combinación de condiciones de causalidad.

Una simple *tabla de verdad* hipotética con cuatro condiciones de causalidad de datos nítidos (es decir, dicotómicas), un resultado y **200** casos se presentan en la **Tabla 8.1.**

**Tabla 8.1. Tabla de verdad hipotética con cuatro *condiciones causales* y una salida: *evitar la disminución de capacidad creativa de la empresa* A**

| Item | Columnas | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | C | I | P | S | A | Número de empresas como casos |
| 1 | 0 | 0 | 0 | 0 | 0 | 30 |
| 2 | 0 | 0 | 0 | 1 | ¿? | 3 |
| 3 | 0 | 0 | 1 | 0 | ¿? | 4 |
| 4 | 0 | 0 | 1 | 1 | ¿? | 0 |
| 5 | 0 | 1 | 0 | 0 | 0 | 25 |
| 6 | 0 | 1 | 0 | 1 | 0 | 19 |
| 7 | 0 | 1 | 1 | 0 | ¿? | 0 |
| 8 | 0 | 1 | 1 | 1 | 0 | 20 |
| 9 | 1 | 0 | 0 | 0 | ¿? | 0 |
| 10 | 1 | 0 | 0 | 1 | ¿? | 1 |
| 11 | 1 | 0 | 1 | 0 | ¿? | 0 |
| 12 | 1 | 0 | 1 | 1 | ¿? | 2 |
| 13 | 1 | 1 | 0 | 0 | ¿? | 19 |
| 14 | 1 | 1 | 0 | 1 | ¿? | 22 |
| 15 | 1 | 1 | 1 | 0 | ¿? | 32 |
| 16 | 1 | 1 | 1 | 1 | ¿? | 23 |
| | | | | | **Total** | **200** |

Fuente: Elaboración propia.

Las cinco *condiciones causales* que producen **A,** son las siguientes:

**a.** El CEO tiene estudios de posgrado, **C.**
**b.** La empresa tiene altos ingresos por innovación, **I.**
**c.** La empresa se dedica a alta tecnología, **P.**

**d.** La empresa tiene un alto puntaje de capacidad innovadora, **S.**
**e.** Evitar la disminución de capacidad creativa de la empresa, **A**.

Explicándose como sigue:

1. Con cuatro *condiciones causales*, hay dieciséis combinaciones lógicamente posibles de condiciones, lo mismo que el número de filas en la tabla. El objetivo del **QCA,** es derivar una declaración lógicamente simplificada, que describa las diferentes combinaciones de condiciones vinculadas a un resultado.
2. Así, el **QCA** presenta la *tabla de verdad* en una forma lógica y resumida. Los datos hipotéticos presentados en la **Tabla 8.1**, muestran una característica de los datos no experimentales: los **200** casos tienen una distribución desigual en las dieciséis filas y algunas combinaciones de *condiciones causales*, en las filas, carecen de casos por completo. El número de empresas, con cada combinación de *condiciones causales,* se reporta en el última columna.
3. En el enfoque de *efectos netos*, esta desigualdad se entiende como resultado de las variables independientes correlacionadas. En general, mientras mayor sea la correlación entre las variables causales, mayor será la desigualdad en la distribución de casos entre las diferentes combinaciones de *condiciones causales*.
4. Por el contrario, en **QCA**, estas desigualdades son entendidas como producto de la diversidad limitada. Desde este punto de vista, las cuatro *condiciones causales* definen los dieciséis tipos diferentes de casos, y las cuatro dicotomías convirtiéndose, por tanto, en una sola variable de escala nominal, con dieciséis posibles categorías.
5. Dado que hay instancias empíricas, de solo un subconjunto de los dieciséis tipos de casos lógicamente posibles, el conjunto de datos es considerado de diversidad limitada.
6. Observe también, que las filas que abarcan solo un pequeño número de casos, no tienen códigos de resultado asignados a ellos (ver *columna 5* **A Tabla 8.1**). Cuando la evidencia empírica es débil, no hay juicios

sobre los resultados de la fila que se puedan hacer. (En el lenguaje de **QCA**, estas filas son *residuales*).

7. En **QCA**, los resultados se codifican utilizando criterios de teoría de conjuntos a partir del marco teórico de la investigación. La pregunta clave, es el grado en que los individuos en la fila, constituyen un subconjunto de las que no están en evitar la *disminución de capacidad creativa de la empresa,* **A.**
8. Eso es para ¿en qué medida los casos en una fila dada coinciden en no desplegar el evitar la disminución de *capacidad creativa de la empresa* **A**? Por supuesto, las relaciones perfectas de subconjuntos son raras con el nivel individual datos. Siempre hay casos sorprendentes, por ejemplo, empresas y CEOs con todas las ventajas posibles, pero que, sin embargo, logran caer en no evitar la *disminución de capacidad creativa de la empresa,* **A.**
9. Con **QCA**, los investigadores establecen reglas para determinar el grado para el cual, los casos en cada fila son consistentes, con la relación del subconjunto. El investigador, primero establece una proporción umbral para la *consistencia* del marco teórico de la investigación, que las proporciones deban exceder.
10. Por ejemplo, un investigador, podría argumentar que la proporción observada de casos en una fila, que no están sobre una fila, en evitar la disminución de *capacidad creativa de la empresa* **A,** debe superar una proporción de referencia de **0.90**. Además, el investigador también puede aplicar criterios probabilísticos convencionales a estas evaluaciones.
11. Por ejemplo, el investigador podría afirmar que la proporción observada de los individuos, que no están en la disminución *de capacidad creativa de la empresa* **A,** deben ser significativamente mayores que un punto de referencia proporción de **0.80**, utilizando un *nivel de significancia* (alfa) de **0.05**. Los puntos de referencia específicos y los alfa utilizados por los investigadores, dependen del estado de los conocimientos sustantivos y teóricos existentes. La evaluación de la *consistencia del marco teórico* de cada fila es directa, cuando las *tablas de verdad* se construyen a partir de conjuntos de datos nítidos.

**12.** Cuando se utilizan *conjuntos difusos*, los principios del marco teórico de la investigación que se invocan, son los mismos, pero los procedimientos son más complejos. La penúltima columna de la **Tabla 8.1** (evitar la *disminución de capacidad creativa de la empresa,* **A**) muestra el codificación del resultado, de cada fila, basado en la evaluación de los puntajes de *consistencia*.

## Comparación de configuraciones

La diferencia clave entre **QCA** y el enfoque de efectos netos, es que este último se centra en variables independientes analíticamente separables y su grado de intercorrelación, mientras que el primero se centra en los tipos de casos definidos, con respecto a las combinaciones de condiciones de causalidad relevantes, que presengtan. En resumen, el enfoque de efectos netos produce generalizaciones, examinando las correlaciones entre variables. **QCA,** construye generalizaciones al comparar casos de configuraciones y construir especificaciones más inclusivas, de las condiciones causales bajo las cuales, una declaración es verdadera. Estas opiniones contrastantes, de la misma evidencia, correlacional vs. configuracional, tienen muy diferentes implicaciones sobre cómo procede el análisis. Observe, por ejemplo, de la **Tabla 8.1** que muestra:

**1.** Una correlación perfecta entre tener el CEO tiene *estudios de posgrado* **(C)** y evitar la disminución de *capacidad creativa de la empresa,* **A.** Es decir, siempre que haya un **1** (sí) en el columna de resultados (evitar la disminución de *capacidad creativa de la empresa*, **A**), también hay un **1** (sí) en el columna *el CEO tiene estudios de posgrado* **C**, así también, siempre que haya un **0** (no) en la evitar la disminución de capacidad creativa de la empresa **A**, también hay un **0** (no) en la columna el CEO *tiene estudios de posgrado,* **C.**

**2.** Desde una perspectiva de *efectos netos*, este patrón constituye una evidencia muy sólida de que la clave para evitar la disminución de *capa-*

*cidad creativa de la empresa* **A,** es asegurar que el CEO tiene *estudios de posgrado,* **C.**

3. Una vez que se tiene en cuenta el efecto de asegurar que el CEO tiene *estudios de posgrado* **C** (utilizando el datos hipotéticos en la **Tabla 8.1**), no queda variación en evitar la disminución de *capacidad creativa de la empresa* **A**, para las otras variables a explicar.
4. Sin embargo, esta conclusión no es tan fácil con **QCA** ya que hay varias combinaciones de *condiciones causales* (es decir, tipos de casos) en la *tabla de verdad*, donde está presente el CEO tiene *estudios de posgrado* **C** y el resultado evitar la disminución de *capacidad creativa de la empresa* **A,** es desconocido, debido a una insuficiencia de casos.
5. Por ejemplo, la **fila 9** combina la presencia el CEO tiene *estudios de posgrado* **C** con ausencia de los otros tres recursos. Sin embargo, no hay casos con esta combinación de condiciones y, en consecuencia, no hay forma de evaluar empíricamente si esta combinación de condiciones está vinculada a evitar la disminución de *capacidad creativa de la empresa,* **A.**
6. Como se explicó en los capítulos previos, en el análisis de *efectos netos* convencionales, tales combinaciones de residuos se incorporan habitualmente en las soluciones; sin embargo, su incorporación es invisible para la mayoría de los investigadores, que usan este enfoque.
7. En el enfoque convencional, los residuos, en efecto, se incorporan encubiertamente a las soluciones mediante la suposición de aditividad: la idea de que el efecto neto de una variable, es el mismo independientemente de los valores de las otras variables independientes. Por lo tanto, el problema de *diversidad limitada* y la *necesidad* de un análisis contrafactual, son ambas consideradas en el esfuerzo analítico, para aislar el efecto de cada variable independiente sobre el resultado.
8. Desde el punto de vista de **QCA**, es notable que las cuatro filas codificadas con **1** (sí) para evitar la disminución de capacidad creativa de *la empresa* **A,** todas incluyen la presencia de ambos: el CEO *tiene estudios de posgrado* **C** y la empresa tiene *altos ingresos por innovación* **I.** Basado en los argumentos presentados en los capítulos previos, no sería

prudente aceptar la conclusión *parsimoniosa*, y asegurar que cuando el CEO tiene estudios de posgrado **C,** por sí sola explica que se evita la disminución de *capacidad creativa de la empresa* **A,** a pesar de que existe una correspondencia perfecta entre ambas, en la **Tabla 8.1**.

**9.** Usando **QCA**, es difícil ignorar el hecho, de que hay dos *condiciones causales*: el CEO tiene estudios de posgrado **C** y la empresa tiene altos ingresos por innovación **I** siempre se combinan para evitar la disminución de capacidad creativa de la empresa **A,** como resultado Estas dos condiciones **(C** e **I)** seguramente se reforzarían entre ellas, en su impacto para obtener **A.**

## Conjuntos difusos y análisis configuracional

Debido a las continuidades matemáticas subyacentes en los conjuntos de datos nítidos y difusos la **Tabla 8.1** podría haberse construido a partir de datos de conjuntos de *datos difusos*. Para hacerlo:

**1.** Habría sido necesario calibrar, el grado de pertenencia de cada caso, en cada uno de los conjuntos definidos por el *condiciones causales* (por ejemplo, el grado de pertenencia en el conjunto de empresas que tiene un alto puntaje de capacidad innovadora, **S.**

**2.** Posteriormente, evaluar el grado de pertenencia de cada caso, en cada una de las dieciséis combinaciones de *condiciones causales* que definen las filas de la **Tabla 8.1**.

**3.** Después de calibrar el grado de pertenencia en el resultado (es decir, en el conjunto de empresas que hacen por evitar la disminución de *capacidad creativa de la empresa,* **A**), sería posible evaluar, el grado de pertenencia en cada combinación de las *condiciones causales* las cuales son un subconjunto difuso, que pertenece al resultado.

**4.** En efecto, estos análisis evalúan el grado en que las empresas, representados en cada fila, hacen por evitar la disminución *de capacidad creativa de la empresa,* **A.**

**5.** Dichas evaluaciones se llevan a cabo, utilizando puntajes de pertenencia difusos, no dicotomizados, y utilizan un criterio más estricto y exigente de definición de la relación de subconjunto, que se utiliza en análisis de conjuntos de datos nítidos.

**6.** Es posible por lo tanto, crear y usar una *tabla de verdad* para resumir el resultados de estas evaluaciones de *datos difusos* En este ejemplo, habría dieciséis evaluaciones de *datos difusos* (porque hay cuatro causales de conjuntos condiciones difusos y, por lo tanto, dieciséis puntajes de pertenencia de configuración).

**7.** Más generalmente, el número de evaluaciones de conjuntos difusos es $\mathbf{2}^{k}$, donde ***k*** es el cantidad de *condiciones causales*. Las filas de la *tabla de verdad* resultante listan las diferentes combinaciones de condiciones evaluadas.

**8.** Por ejemplo, la **fila 4** de la *tabla de verdad* resumiría los resultados del análisis de *datos difusos* del grado de pertenencia en el conjunto de empresas, que combinan baja pertenencia en el CEO tiene *estudios de posgrado,* **C;** baja pertenencia a la empresa tiene altos *ingresos por innovación,* **I;** alta pertenencia de la *empresa se dedica a alta tecnología,* **P** y alta pertenencia a la empresa tiene un *alto puntaje de capacidad innovadora,* **S.**

**9.** Tenga en cuenta que con *conjuntos difusos*, el problema de la *diversidad limitada* se transforma de una de las celdas vacías, en una tabulación cruzada (*k-way cross-tabulation*) de *condiciones causales* dicotomizadas (es decir, las filas restantes en una *tabla de verdad*), a uno de los sectores vacantes (o mayormente vacantes) en un espacio vectorial con *k* dimensiones difusas.

**10.** Los $\mathbf{2}^{k}$ sectores de este espacio, varían en el grado de que se rellenan con casos, y algunos sectores carecen de casos en total. En otras palabras, con datos sociales naturales, es común para muchos sectores del espacio vectorial definidos por *condiciones causales*, ser nulo en casos, así como es común, para una tabulación cruzada (*k-way cross-tabulation*) de dicotomías, producir una abundancia de celdas vacías.

**11.** Las mismas herramientas desarrolladas para abordar la diversidad limitada en análisis de conjuntos de datos nítidos, se pueden utilizar para abordar la diversidad limitada en análisis de conjuntos de *datos difusos*, utilizando la *tabla de verdad* que resume los resultados del análisis del conjunto de *datos difusos*, como un paso intermediario.

**12.** La columna de resultados, en una *tabla de verdad* de un conjunto de datos difuso, muestra los resultados de las $\mathbf{2}^{k}$ evaluaciones, es decir, si el grado de pertenencia está o no en la configuración de *condiciones causales*, especificadas en una fila, pueden ser considerados un subconjunto difuso del grado de pertenencia en el resultado.

**13.** El examen de la *tabla de verdad* resultante es, en efecto, un análisis de declaraciones que resumen los $\mathbf{2}^{k}$ análisis de conjuntos de *datos difusos*. El producto final, el análisis de la *tabla de verdad*, a su vez, una síntesis lógica de estas afirmaciones. Esta síntesis, especifica las diferentes combinaciones de condiciones de causalidad vinculadas al resultado, a través de la relación de subconjunto de datos difuso.

El capítulo 9, ofrece una demostración completa del uso de conjuntos difusos en un caso hipotético de investigación social orientada a la innovación, utilizando un conjunto de muestras con **N** grande. Contrasta dos análisis de los mismos datos, un análisis convencional de *efectos netos* usando regresión logística y un análisis del conjunto de *datos difusos* utilizando el enfoque de *tabla de verdad* recién descrito. El contraste entre estos dos análisis muestra que el enfoque de conjunto de *datos difusos*, ofrece lo que se omite en un análisis de *efectos netos*, a saber, por lo que se sugiere prestar mucha atención a casos como *configuraciones* y a las diferentes combinaciones de condiciones que están vinculados a un resultado.

CAPÍTULO 9.

# Efectos netos vs. Configuraciones en fsQCA

Este capítulo, presenta una crítica a la corriente de *efectos netos* en una manera práctica, contrastando un análisis convencional de *efectos netos* con una cantidad de muestras **N** grande de datos provenientes de la **Tabla 8.1** conjunto de datos relevantes para evitar la *disminución de capacidad creativa de la empresa* **A**, con un análisis alternativo de los mismos datos, siguiendo los principios desarrollados en este libro. Mientras que la dos enfoques, difieren en varios aspectos importantes, la diferencia clave es que:

**a.** El *enfoque de efectos netos*, se centra en los efectos independientes de variables causales en el resultado, mientras que

**b.** El *enfoque configuracional*, atiende a combinaciones de *condiciones causales* e intenta establecer enlaces explícitos entre combinaciones específicas de condiciones y el resultado. Este método alternativo, cono-cido como análisis *cualitativo comparativo de conjunto difuso* (**fsQCA**), combina el uso de conjuntos de *datos difusos* con el análisis de casos como *configuraciones*, una característica central de la investigación social orientada a casos **(Ragin, 1987)**. En este enfoque, cada caso se examina en términos del grado de pertenencia, en diferentes combina-ciones de condiciones de causalidad relevantes.

Usando **fsQCA**, los investigadores pueden considerar que los casos varían por grado de pertenencia, en todas las combinaciones lógicamente posibles, de un conjunto dado de *condiciones causales* y posteriormente usar métodos para analizar el marco teórico de la investigación, de manera lógi-camente disciplinada, las variadas conexiones de combinaciones causales y el resultado. Se ofrece este enfoque alternativo, no como un reemplazo del análisis de *efectos netos*, sino como una técnica complementaria. El

**fsQCA,** se entiende mejor como una técnica de exploración / interpretación, basada en la teoría de conjuntos. Si bien, los criterios probabilísticos pueden ser incorporados en el **fsQCA**, no es una técnica inferencial, *per se*. Es una forma alternativa de analizar la evidencia, a partir de supuestos muy diferentes, con respecto a los tipos de hallazgos que buscan los científicos sociales orientados a la innovación. Estos supuestos alternativos, reflejan la lógica y el espíritu de la investigación cualitativa, donde los investigadores estudian casos como *configuraciones*, interesados en como se ajustan las diferentes partes o aspectos de los casos.

## Caso hipotético con el análisis de efectos netos

Un análisis de *efectos netos* de los datos, se realiza en el modelo hipotético, mencionado en la **Tabla 8.1**, y con datos complementarios a saber:

**a.** *El CEO tiene estudios de posgrado,* **C.**
**b.** *La empresa tiene altos ingresos por innovación,* **I.**
**c.** *La empresa se dedica a alta tecnología,* **P.**
**d.** *La empresa tiene un alto puntaje de capacidad innovadora,* **S.**
**e.** *Evitar la disminución de capacidad creativa de la empresa,* **A.**

Incluyendo:

**f.** *Edad del CEO encuestado,* **E.**
**g.** *Posicionamiento en el mercado,* **M.**
**h.** *Hace productos considerando el impacto social,* **B.**

Calculado con un análisis básico de regresión logística, para medir la importancia de la puntuación de **S** en una variedad de resultados. Son controlados los efectos de dos variables, a saber **S** y **E.** Un hallazgo principal es que se encontró que **S** es más importante que **E** respecto a la salida **A.** Esto significa que la empresa al tener un alto puntaje de *capacidad*

*innovadora* **S**, se encuentra más comprometida a evitar la *disminución de capacidad creativa de la empresa* **A**, sin depender de la edad del CEO **E** de la misma.

**Tabla 8.1. Tabla de verdad hipotética con cuatro *condiciones causales* y una salida: *evitar la disminución de capacidad creativa de la empresa A*.**

| Item | Columnas | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| | C | I | P | S | A | Número de empresas como casos |
| 1 | 0 | 0 | 0 | 0 | 0 | 30 |
| 2 | 0 | 0 | 0 | 1 | ¿? | 3 |
| 3 | 0 | 0 | 1 | 0 | ¿? | 4 |
| 4 | 0 | 0 | 1 | 1 | ¿? | 0 |
| 5 | 0 | 1 | 0 | 0 | 0 | 25 |
| 6 | 0 | 1 | 0 | 1 | 0 | 19 |
| 7 | 0 | 1 | 1 | 0 | ¿? | 0 |
| 8 | 0 | 1 | 1 | 1 | 0 | 20 |
| 9 | 1 | 0 | 0 | 0 | ¿? | 0 |
| 10 | 1 | 0 | 0 | 1 | ¿? | 1 |
| 11 | 1 | 0 | 1 | 0 | ¿? | 0 |
| 12 | 1 | 0 | 1 | 1 | ¿? | 2 |
| 13 | 1 | 1 | 0 | 0 | ¿? | 19 |
| 14 | 1 | 1 | 0 | 1 | ¿? | 22 |
| 15 | 1 | 1 | 1 | 0 | ¿? | 32 |
| 16 | 1 | 1 | 1 | 1 | ¿? | 23 |
| | | | | | **Total** | **200** |

Fuente: Elaboración propia.

Los principales hallazgos en el modelo con la presencia / ausencia de evitar la disminución de *capacidad creativa de la empresa* **A**, como resul-tado de interés, se presenta en la**Tabla 9.1** (con ausencia **A = 1**).

**Tabla 9.1. Regresión logística para evitar la disminución de capacidad creativa de la empresa A con el puntaje de S , M y E**

| Descripción | B | S.E. | Sig. | Exp (B) |
|---|---|---|---|---|
| La empresa tiene un alto puntaje de *capacidad innovadora,* **S** (z score) | 0.651 | 0.139 | 0.000 | 1.917 |
| *Posicionamiento en el mercado,* **M** (z score) | 0.376 | 0.117 | 0.001 | 1.457 |
| Edad del CEO encuestado **E** | 0.040 | 0.050 | 0.630 | 1.040 |
| Constant | 1.123 | 0.859 | 0.191 | 3.074 |

Nota: Chi-squared= 53.973, df= 3. B= regression coefficient, S.E.=Standard error of regression coefficient, Sig.= statistical significance

Fuente: Elaboración propia.

El análisis reportado, utiliza datos estandarizados **(*z score*)** para ambos casos **S** y **M** para facilitar la comparación de efectos. El análisis que se muestra en la **Tabla 8.1,** se limita a CEOs provenientes de universidades públicas, con datos completos sobre todas las variables utilizadas en este análisis y en los posteriores informados en este capítulo, incluido el análisis de *datos difusos*. El fuerte impacto positivo de los puntajes **S**, a pesar de las estadísticas control para el efecto de **M**, refleja una curva de campana en los resultados.

La tesis anterior, es posible rebatirla si a partir de una regresión logística más elaborada, tomando en cuenta los datos anteriores, donde paso a paso, se podrían incluir más y más *condiciones causales* (por ejemplo, **C, I, P**) argumentando que deberían verse como factores que intervienen sobre **A**. Una opinión, es que **S** tiene un efecto sustancial en el análisis de la curva de campana, porque en la regresión logística estuviera poco descrita. Para remediar este problema, sería factible, por marco teórico y empírico sustantivo del investigador, aumentar más factores (al menos quince) **S** en la posibilidad de que influyan sobre el resultado **A**. Este enfoque, tiene también la posibilidad de que se reduzca dramáticamente el impacto de los puntajes de **S** sobre **A**, circunstancia que deja abierta la duda razonada, de una especificación mal planeada al ser sobre-inclusiva.

La **Tabla 9.2** informa los resultados de un análisis de regresión logística de evitar la disminución de *capacidad creativa de la empresa* **(A)** usando solo un número moderado de variables independientes.

**Tabla 9.2. Regresión logística para evitar la disminución de capacidad crea-tiva de la empresa A con el puntaje de S, I, C, P y B**

| Descripción | B | S.E. | Sig. | Exp (B) |
|---|---|---|---|---|
| La empresa tiene un alto puntaje de *capacidad innovadora,* **S** (z score) | 0.391 | 0.154 | 0.011 | 1.479 |
| La empresa tiene altos *ingresos por innovación*, **I** (z score) | 0.357 | 0.154 | 0.020 | 1.429 |
| El CEO tiene *estudios de posgrado*, **C** (z score) | 0.635 | 0.139 | 0.000 | 1.887 |
| La empresa se dedica a *alta tecnología*, **P** (Sí=1; 0=No) | 1.658 | 0.346 | 0.000 | 5.251 |
| Hace productos considerando el *impacto social*, **B** (Sí=1; 0=No) | -0.524 | 0.282 | 0.063 | 0.592 |
| Constant | 1.970 | 0.880 | 0.025 | 7.173 |

Nota: Chi-squared= 104.729, df= 5. B= regression coefficient, S.E.=Standard error of regression coefficient, Sig.= statistical significance

Fuente: Elaboración propia.

Específicamente, la presencia / ausencia de **A** (con ausencia = 1) es producto de la intervención de cinco variables independientes **S, I, C, P, B**. Las tres variables de escala de intervalos, están estandarizadas (usando puntajes z score) para simplificar la comparación de efectos. Como en el anterior análisis, la tabla muestra los resultados solo para CEOs provenientes de universidades públicas. La razón de esta especificación, es la de comparar la **Tabla 9.1** de dos variables independientes vs. la **Tabla 9.2** que usa cinco variables independientes y observar que el análisis de regresión logística intenta encontrar un equilibrio entre las dos extermos de las especificaciones, mientras se enfoca en varios de las más importantes *condiciones causales*. Por lo tanto, se tiene:

1. Los resultados presentados en la **Tabla 9.2 (B=0.391)** son consistentes con los resultados de la **Tabla 9.1** ya que ambos muestran que la pun-

tuación de **S** tiene un impacto independiente en la prevención **A**, pero no tan fuerte como el informado por **Tabla 9.1** (**B=0.651**).

2. La **Tabla 9.2** muestra efectos muy fuertes de las *condiciones causales* contribuyentes, especialmente en (**C**, **B=0.635**) y (**P**, **B=1.658**)
3. La **Tabla 9.2** confirma la especificación de dependencia del análisis de *efectos netos*. Por ejemplo, si **C** se acepta como un causal contribuyente (y no se considera derivado de los puntajes **S**), entonces es claramente más importante que los puntajes de **S**.
4. Asimismo, el *impacto* **P** en las posibilidades de mantener fuera **A** es sustancial para los CEOs provenientes de universidades públicas.
5. Según la **Tabla 9.2**, los CEOs provenientes de universidades públicas con la empresa se dedica a alta tecnología (**P**, **Exp (B) = 5.251**) tienen cinco veces más probabilidades de lograr **A** que los CEOs provenientes de universidades públicas con no tener la empresa que se dedica a *alta tecnología,* **P.**
6. Incluso aunque estos resultados, como todos los análisis de *efectos netos*, dependen de las especificaciones, el hecho de que un número muy modesto de variables independientes, reduzcan en gran medida la estimación del efecto de las puntuaciones **S** sobre la **A**, puede arrojar dudas sustanciales sobre su grado de pertenencia en el resultado.

## Caso hipotético con fsQCA

El éxito de cualquier análisis de conjuntos *difusos* depende tanto de la construcción cuidadosa como de su *calibración*. Lo principal de los tipos de análisis, tanto de *datos nítidos* como *datos difusos*, es la evaluación de las relaciones del marco teórico de la investigación, por ejemplo, la evaluación de si la pertenencia, en una combinación de las *condiciones causales*, puedan considerarse, un subconjunto consistente de la pertenencia en un resultado dado. Se dice que existe una relación de subconjunto de *datos difusos*, cuando las puntuaciones en un solo conjunto de relaciones causales (por ejemplo, el conjunto de *datos difusos*, de empresas que combinan

**C, P, I, S, M, E, B** (mencionados en la **Tabla 9.2**) son consistentemente menores o igual que los puntajes del conjunto opuesto (por ejemplo, el conjunto difuso de individuos que no están en la **A**. Por lo tanto, importa mucho cómo se construyen los conjuntos de *datos difusos* y cómo se calibran los puntajes de pertenencia. Las calibraciones incorrectas graves, pueden distorsionar o socavar la identificación de las relaciones del marco teórico de la investigación. Por el contrario, para que la variable convencional sea útil en un análisis de *efectos netos*, solo necesita variar de manera significativa.

A menudo, la métrica específica de una variable convencional, es ignorada por los investigadores por completo porque es arbitraria o sin sentido. Incluso cuando una variable tiene una métrica significativa, los investigadores a menudo se centran solo en la dirección y la importancia de su efecto. Para calibrar los puntuaciones de pertenencia del conjunto de datos difuso, los investigadores deben usar su conocimiento empírico sustantivo. Los puntajes de pertenencia resultantes, deben tener validez aparente (*face validity*) en relación con el conjunto de relaciones en cuestión, especialmente sobre cómo está conceptualizado y etiquetado. Una puntuación de datos difusa de **0.25**, por ejemplo, tiene un significado muy específico. Esto es, si un caso está a medio trayecto entre exclusión total de un conjunto, por ejemplo, un puntaje de pertenencia de **0.0** de, **I**) y el punto de cruce o traslape (*crossover*) de **0.5**, (el punto de máxima ambigüedad en si, un caso está más dentro o más fuera del conjunto). Como se explica en **Ragin (2000)**, las decisiones más importantes en la calibración de un conjunto de datos difuso, implican la definición de tres referencias cualitativas: inclusión completa en el conjunto, el punto de cruce (pertenencia = **0.5**) y exclusión total de el conjunto.

Los conjuntos de relaciones principales, utilizados en el análisis reportado en este capítulo, son el conjunto de empresas que aplican **A** y el grado de pertenencia en conjuntos de relaciones que reflejen cinco puntuaciones antecedentes de los factores**: S, I, C, P** y **B**. La calibración de estos conjuntos de *datos difusos*, se detalla al final de este capítulo.

En este punto, es importante tener en cuenta, que representar una sola variable de escala de intervalo, con dos conjuntos de *datos difusos* suele ser muy útil. Por ejemplo, la variable la empresa tiene altos *ingresos por innovación* **I** puede ser transformada separadamente dentro del conjunto de empresas con bajos **I**. Es necesario construir dos conjuntos de *datos difusos*, debido a la asimetría de los dos conceptos objetivo. La no pertenencia completa en el conjunto de empresas con *ingresos por innovación altos* **I** (una puntuación de pertenencia de **0.0**) no implica pertenencia completa en el conjunto de empresas con *ingreso por innovación bajos* **I** (una puntuación de pertenencia de **1.0**), porque es posible estar completamente fuera del conjunto de empresas con *ingresos por innovación altos* **I** sin estar completamente en el conjunto de empresas con bajos **I.** Lo mismo es cierto, para las otras dos variables de escalas de intervalos utilizadas como *condiciones causales* en el análisis de la regresión logística (**Tabla 9.2**), con las puntuaciones de **S** y **C**. Esta codificación dual de las *condiciones causales* clave, tiene un importante beneficio teórico. Por ejemplo, ¿se tiene un alto puntaje **S** que esté vinculado a una **A** con mayores posibilidades, o no se tiene un bajo puntaje **S** que importe? Este problema es especialmente importante, porque tener un puntaje **S** alto como la **Tabla 9.1** lo indica, significa que es la clave del éxito, en las empresas de hoy en día. Tenga en cuenta también, el lenguaje y la lógica de las variables, no se traducen directamente del marco teórico de la investigación. Por ejemplo:

1. Un caso que no puede tener pertenencia en una variable, como un alto grado de pertenencia en **S** o en **I**.
2. En cambio, un caso puede tener pertenencia en un conjunto, por ejemplo, pertenencia muy alta en el conjunto de empresas con puntajes altos de **S** o en **I.**

La traducción de variables a conjuntos requiere cuidado definición y etiquetado de los objetivos de los conjuntos de relaciones causales, que a su vez, proporciona la base principal para calibrar la pertenencia. Por lo tanto, al traducir variables como **I** a *conjuntos difusos*, es útil considerar

los diferentes los objetivos de los conjuntos de relaciones causales que se pueden crear, a partir de una sola variable de origen, especialmente a la luz de las cuestiones teóricas y empírico sustantivas que inspiran y guian la investigación.

En total, el análisis del conjunto de *datos difusos* hipotético, que se reporta en este capítulo, utiliza ocho *condiciones causales*:

1. Dos, son de conjuntos de relaciones nítidos: **P** vs. **(~P)** y **(B)** y **(~B).**
   **a.** La empresa se dedica a alta tecnología (**P**, Sí=**1**; **~P,** No=**0**)
   **b.** Hace productos considerando el impacto social (**B**, Sí=**1**; **~B,** No=**0**)
2. Los seis restantes, son conjuntos de *datos difusos* con un grado de pertenencia en el conjunto de casos, con:
   **a.** *La empresa tiene un alto puntaje de capacidad innovadora,* **S.**
   **b.** *La empresa tiene un bajo puntaje de capacidad innovadora,* **(~S).**
   **c.** *La empresa tiene altos ingresos por innovación,* **I.**
   **d.** *La empresa tiene bajos ingresos por innovación,* **(~I).**
   **e.** *El CEO tiene estudios de posgrado,* **C**
   **f.** *El CEO tiene estudios de pregrado,* **(~C).**
3. Después de calibrar el conjunto de *datos difusos*, la siguiente tarea es calcular el grado de pertenencia de cada caso, en cada uno de los $\mathbf{2}^{k}$, combinaciones lógicas posibles, de ocho *condiciones causales*,
4. Posteriormente, evaluar la distribución de casos en estas combinaciones. Con ocho *condiciones causales*, hay **256** combinaciones de condiciones lógicamente posibles. Por supuesto, muchas de estas **256** combinaciones no son empíricamente posibles. Por ejemplo, un caso no puede tener una alta pertenencia tanto en la empresa tiene *altos ingresos por innovación* **I** como en la empresa no tiene *altos ingresos por innovación* **I.** El número de combinaciones empíricamente posibles es **108**. Este número todavía es de considera contra el número de combinaciones de alta frecuencia (ver **Tabla 9.3**).

**Tabla 9.3 Distribución de casos através del vector de espacio por esquinas (42 combinaciones con al menos 4 casos cada uno)**

| Item | Columnas | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | P | B | I | | S | | C | | Count | Cumulative proportion |
| | | | Alta | Baja | Alta | Baja | Pregrado | Posgrado | | |
| 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 118 | 0.152 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 78 | 0.253 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 53 | 0.321 |
| 4 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 41 | 0.375 |
| 5 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 39 | 0.425 |
| 6 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 34 | 0.469 |
| 7 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 30 | 0.508 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 23 | 0.537 |
| 9 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 22 | 0.566 |
| 10 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 20 | 0.592 |
| 11 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 20 | 0.618 |
| 12 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 19 | 0.642 |
| 13 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 19 | 0.667 |
| 14 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 18 | 0.690 |
| 15 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 12 | 0.705 |
| 16 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 12 | 0.721 |
| 17 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 11 | 0.735 |
| 18 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 11 | 0.749 |
| 19 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 11 | 0.764 |
| 20 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 11 | 0.778 |
| 21 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 10 | 0.791 |
| 22 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 10 | 0.804 |
| 23 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 10 | 0.817 |
| 24 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 9 | 0.828 |
| 25 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 9 | 0.840 |
| 26 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 9 | 0.851 |
| 27 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 7 | 0.860 |
| 28 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 7 | 0.870 |
| 29 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 6 | 0.877 |
| 30 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 6 | 0.885 |
| 31 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 6 | 0.893 |
| 32 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 6 | 0.901 |
| 33 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 6 | 0.908 |
| 34 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 5 | 0.915 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 35 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 5 | 0.921 |
| 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0.926 |
| 37 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 4 | 0.932 |
| 38 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 4 | 0.937 |
| 39 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 4 | 0.942 |
| 40 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 4 | 0.947 |
| 41 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 4 | 0.952 |
| 42 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 4 | 0.957 |
| **All remaining combinations of conditions; row counts** | | | | | | | | | **< 4** | **1.000** |

Fuente: Elaboración propia.

5. La **Tabla 9.3**, lista **42 de 256** combinaciones que tienen al menos cuatro casos con más de **0.5** miembros.Un adicional de **19** filas (no mostradas en la **Tabla 9.3**, tienen uno, dos o tres casos cada uno. El resto de las filas, no tienen con pertenencias mayores a **0.5**.
6. Recuerde, que un caso puede tener, a lo sumo, solamente un puntaje de pertenencia qmayor que **0.5**. Por lo tanto, las **256** combinaciones de condiciones se pueden evaluar, con respecto a la frecuencia de casos examinando el número de instancias empíricas de cada combinación.
7. Si una configuración no tiene casos, mayores a **0.5** de pertenencia, entonces no hay casos que estén más dentro que fuera del conjunto, definido por la combinación de condiciones (y sin casos en el correspondiente sector del vector de espacio multidimensional, definido por las *condiciones causales*).
8. La **Tabla 9.3** revela, que los datos utilizados en este análisis (y, por implicación, en los análisis de regresión logística reportados en las **Tablas 9.1 y 9.2**) están notablemente limitados en su diversidad. En total, solo **42 de los 256** sectores contenidos dentro del vector de espacio de ocho dimensiones, tienen al menos cuatro instancias empíricas de **0.5** de pertenencia en la esquina) y, la mayoría de las frecuencias reportadas en la tabla son bastante pequeños.
9. Los dos sectores más poblados, representan el **25%** de los casos; las siete capturas más pobladas, la mitad de los casos; y, las catorce capturas más pobladas casi **70%** de los casos. El número de sectores bien poblados

(catorce) es más bien pequeño, incluso en relación con el número de sectores que existen en cinco dimensiones del espacio vectorial (**$2^k$= 32 *k*= 5**).

10. Este es el número de sectores que deberían haberse obtenido, si las tres variables de intervalo **C, I** y **S** utilizados en el análisis de regresión logística habrían sido transformadas, dentro en un conjunto de datos difuso, en lugar de dos.
11. En el análisis de *datos difusos* de este tipo (con cantidad grande de muestras **N**), es importante establecer un fuerte umbral de evidencia, para las combinaciones de *condiciones causales*, utilizando la información sobre la distribución de fuertes instancias a través de los sectores.
12. Específicamente, las combinaciones causales con solo unas pocas instancias fuertes (es decir, algunos casos con más de **0.5** de pertenencia en la combinación) deben filtrarse y no estar sujeto a más análsis empíricos. Además del hecho, de que sería imprudente basar una conclusión sobre una combinación de atributos a nivel individual, en un pequeño número de instancias, la existencia de casos en sectores de baja frecuencia puede deberse a un error de medición o asignación.
13. El análisis de *datos difusos* que sigue, utiliza un umbral de frecuencia de al menos diez instancias fuertes. Este valor se seleccionó porque captura más del **80%** de los casos asignados a las combinaciones causales.
14. Usando esta regla, las veintitrés combinaciones de condiciones más comunes se mantienen en este análisis. Las filas de baja frecuencia (incluidas las que se muestran en la parte inferior parte de la **Tabla 9.3,** con frecuencias que van de cuatro a nueve) se filtran del análisis. Debido a que estas filas no cumplen con el umbral de solidez de la evidencia, se tratan como combinaciones *residuales* en el análisis que sigue.
15. La siguiente tarea, es evaluar la *consistencia* de la evidencia para cada una de las combinaciones de *condiciones causales* (las veintitrés frecuencias altas filas de la **Tabla 9.3**) con la relación de subconjunto. Específicamente, es necesario determinar si el grado de pertenencia en cada combinación de *condiciones causales*, es un subconjunto del grado de pertenencia en el resultado.

**16.** Como se explicó en capítulos previos, la relación de subconjunto se usa para evaluar el ser suficiente de la relación causal. Con conjuntos de *datos difusos*, la relación de subconjunto se demuestra mediante probar que ese grado de pertenencia, en una combinación de condiciones (que puede variar de **0.0 a 1.0**), es consistentemente menor o igual que el grado de pertenencia en el resultado. Estas evaluaciones usan todos los casos en cada evaluación, incluidos los casos con menos de **0.5** de pertenencia en una combinación dada.

**17.** Tales casos pueden ser inconsistentes dependiendo de la relación del marco teórico de la investigación. Por ejemplo, un caso con una pertenencia de **0.40** en una combinación causal y una la pertenencia de **0.20** en el resultado, reduciría el puntaje de *consistencia* para esa combinación, a pesar de que este caso está más fuera que dentro para ambos, la combinación y el resultado.

**18.** Como se muestra en el capítulo 4, una medida simple y descriptiva del grado en el cual, la evidencia con respecto a una combinación de condiciones es consistente, con la relación de subconjunto con respecto al resultado, es:

$$\Sigma\ [\min\ ( X_i\ ,\ Y_i)]\ /\ \Sigma\ ( X_i)$$

**19.** Donde **min** indica la selección del menor de las dos puntuaciones, $X_i$ indica el grado de pertenencia en una combinación de condiciones, y $Y_i$ indica el grado de pertenencia en el resultado. Cuando todos los valores de $X_i$ son consistentes (es decir, sus puntajes de pertenencia en la combinación, son uniformemente menores o iguales que sus valores de correspondientes), el cálculo produce una puntuación de **1.0.**

**20.** Si muchos de los valores de $X_i$, exceden sus valores de $Y_i$ por un margen sustancial, sin embargo, la puntuación resultante es sustancialmente menos de **1.0.**

**21.** En general, los puntajes en esta medida que son inferiores a **0.75,** indican una desviación sustancial de la relación marco teórico de la investigación $X_i <= Y_i$.

**22.** La **Tabla 9.4** informa los resultados de las evaluaciones de *consistencia* del marco teórico de la investigación para las veintitrés combinaciones en la **Tabla 9.3** que cumplen con el umbral de solidez de la evidencia (una frecuencia de al menos diez casos que están más dentro que fuera de cada combinación).

**Tabla 9.4 Evaluacion de la consistencia del marco teórico de la investigación para 23 configuraciones con frecuencia de umbral en los últmos 10 casos**

| Item | Columnas | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| | P | B | I | | S | | C | | Count | Consis-tency | Out-come |
| | | | Alta | Baja | Alta | Baja | Pregrado | Posgrado | | | |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 12 | 0.986 | 1 |
| 2 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 10 | 0.893 | 1 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 12 | 0.892 | 1 |
| 4 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 11 | 0.884 | 1 |
| 5 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 10 | 0.876 | 1 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 23 | 0.864 | 1 |
| 7 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 19 | 0.830 | 1 |
| 8 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 11 | 0.792 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 53 | 0.788 | 0 |
| 10 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 10 | 0.767 | 0 |
| 11 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 39 | 0.754 | 0 |
| 12 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 41 | 0.706 | 0 |
| 13 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 34 | 0.657 | 0 |
| 14 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 19 | 0.641 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 78 | 0.636 | 0 |
| 16 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 20 | 0.620 | 0 |
| 17 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 30 | 0.617 | 0 |
| 18 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 11 | 0.578 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 11 | 0.498 | 0 |
| 20 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 118 | 0.482 | 0 |
| 21 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 22 | 0.402 | 0 |
| 22 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 20 | 0.376 | 0 |
| 23 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 18 | 0.340 | 0 |

Fuente: Elaboración propia.

**23.** Los puntajes de *consistencia*, para las combinaciones causales, varían en un rango de **0.340** a **0.986**, lo que indica un rango considerable de las puntuaciones que satisfacen la relación del subconjunto.

**24.** En el análisis de la tabla verdad, las siete combinaciones con puntuaciones de *consistencia* de al menos **0.80** se tratan como subconjuntos del resultado; los restantes dieciséis, no cumplen este criterio. Una vez que se hace esta distinción, la **Tabla 9.4** puede analizarse como una *tabla de verdad*.

**25.** El resultado expresado en binario **(1,0),** que se basa en las puntuaciones de *consistencia* del marco teórico como conjunto de datos difuso, aparece en la última columna de la **Tabla 9.4**. Con el uso de **fsQCA (Ragin, et al., 2007)**, es posible generar dos soluciones de *tabla de verdad*, una con maximización *parsimoniosa* y la otra, con maximización de la *complejidad*. La solución más *parsimoniosa*, permite la incorporación de cualquier combinación contrafactual que contribuya a la generación de una solución lógicamente más simple.

**26.** Esta solución de la *tabla de verdad* produce tres combinaciones rela- tivamente simples, vinculado a a la evitar la disminución de *capacidad creativa de la empresa*, **A:**

**(P) * ( ~B) + (I) * (~S) + (~C) * (~S)**

**27.** La solución *parsimoniosa*, revela que las tres combinaciones de *condiciones causales* vinculadas a **A**, son:

- La empresa se dedica a *alta tecnología* **P**, combinado con no hace productos considerando el impacto social **(~B).**
- La empresa tiene altos *ingresos por innovación* **I**, combinados con la empresa tiene un bajo puntaje de capacidad innovadora **(~S),** y
- El CEO tiene estudios de pregrado **(~C)** combinado la empresa no tiene un bajo puntaje de capacidad innovadora **(~S).**

**28.** Aunque *parsimoniosa*, esta solución incorpora muchas combinaciones *contrafactuales* (es decir, *residuales*), y muchas de ellas, a su vez, son *contrafactuales difíciles* desde la perspectiva del conocimieto existente, teórico y empírico sustantivo.

**29.** Por ejemplo, la combinación el CEO tiene estudios de pregrado **(~C)** combinado con la empresa se dedica a alta tecnología **P** y con *no hace productos considerando el impacto social* **(~B)** está incluido en las combinaciones mencionadas anteriormente. Muy pocos casos empíricos de esta combinación, están presentes para permitir su evaluación, pero la solución *parsimoniosa* supone que los casos con esta combinación, son capaces de evitar la disminución de capacidad creativa de la empresa **A,** a pesar de su fracaso para lograr que que el CEO *tiene estudios de pregrado* **(~C).**

**30.** Con 256 combinaciones de *condiciones causales* lógicamente posibles, muchas las combinaciones se presentan sin o muy pocos casos, como lo indica la **Tabla 9.3**. La solución *parsimoniosa* que se acaba de presentar, incorpora muchas de tales combinaciones, sin tener en cuenta su factibilidad empírica es decir, sin tener en cuenta el conocimiento empírico sustantivo existente.

**31.** Si, en cambio, el investigador evalúa la factibilidad de las combinaciones *contrafactuales*, una solución menos *parsimoniosa* (intermedia) se puede derivar.

**32.** El paquete de software **fsQCA,** es capaz de producir las tres soluciones (compleja, *parsimoniosa*, e intermedio) cuando se hace clic en el botón Análisis estándar (*Standard Analyses*) en la parte inferior de la hoja de cálculo de la *tabla de verdad*. Posteriormente, se le solicita al usuario el ingreso de datos, como base para la generación de la solución intermedia.

**33.** Esta solución intermedia, se obtiene generando primero la solución más compleja (no se muestra aquí) y posteriormente usar solo el contrafactual fácil para producirla. El conocimiento empírico sustantivo, que se incorpora a la generación de la solución intermedia, en el presente análisis, es bastante simple. Por ejemplo, haga una cadena de suposiciones como: que el CEO *tiene estudios de posgrado* **C** (en lugar de pregrado **~C**) está vinculado a evitar la disminución de *capacidad creativa de la empresa* **A**, que la empresa tiene *bajos ingresos por innovación* **(~I)**

está vinculado a **A** que *la empresa se dedica a alta* tecnología **P** está vinculado a **A**, y así.

**34.** La solución intermedia, es un subconjunto de la solución más *parsimoniosa* y un superconjunto de la solución más compleja. La solución intermedia indica que cinco combinaciones de condiciones están vinculados a evitar la disminución de *capacidad creativa de la empresa*, **A:**

**(P) * ( ~B) * (C) + (P) * (I) * (~ S)* (C) + ( ~B) * (I) * (~ S) * (C) +**
**( ~B) * (~ S) * (~C) + (P) * ( ~ I) * (~ S) * (~C)**

**35.** Estas cinco combinaciones relacionadas con evitar la disminución de capacidad creativa de la empresa **A**, tienen:

**a.** Es común que el CEO tiene estudios (de pregrado **C** o posgrado **~ C)** y algún aspecto relacionado con la empresa (dedica a *alta tecnología* , **P** o hace *productos considerando el impacto social* **,~B**).

**b.** Cuatro incluyen la empresa no tiene un bajo puntaje *de capacidad innovadora* **(~S)** y tres incluyen condiciones varias (altas/bajas) relacionadas con *ingresos por innovación,* **I.**

**36.** Estos resultados son importantes, porque confirman que las *condiciones causales* vinculadas a evitar la disminución *de capacidad creativa de la empresa* **A**, es de naturaleza combinatoria y es posible discernir las combinaciones relevantes, cuando los casos se ven como *configuraciones*.

**37.** Recordemos, que los términos que conforman la solución *parsimoniosa*, deben incluirse en cualquier representación de los resultados, para esto, son decisivos los componentes causales que distinguen las combinaciones de condiciones que son subconjuntos consistentes de los resultados, de aquellos que no lo son (es decir, entre las combinaciones que pasan la frecuencia límite). Por lo tanto, estos factores deben considerarse el núcleo de las *condiciones causales*. Los factores que se agregan en la solución intermedia, son aquellos que también están presentes en los casos que, consistentemente muestran el resultado, pero que requieren *contrafactuales difíciles* para eliminar.

**38.** Por lo tanto, estas *condiciones causales*, son condiciones complementarias o contribuyentes, ya que tienen sentido como factores de contribución importantes y pueden eliminarse de la solución, solo si el investigador está dispuesto a hacer suposiciones que estén en desacuerdo con el conocimiento empírico sustantivo y el marco teórico de la investigación.

**39.** La **Tabla 9.5** resume las cinco soluciones de una manera que diferencía las *condiciones causales* centrales vs. complementarias.

**Tabla 9.5. Condiciones para lograr la evitar la disminución de capacidad creativa de la empresa A, de acuerdo a la educación CEOs provenientes de universidades públicas**

| | Solution | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| **Estatus de empresa** | | | | | |
| Se dedica alta tecnología (P) | ● | • | | | • |
| Hace productos de impacto social (B) | ⊖ | | ⊝ | ⊝ | |
| **Educación del CEO (C)** | | | | | |
| Pregrado | • | • | • | | |
| Posgrado | | | | ● | ● |
| **Capacidad innovadora (S)** | | | | | |
| Alta | | | | | |
| Baja | | ⊖ | ⊖ | ⊖ | ⊖ |
| **Ingresos por innovación (I)** | | | | | |
| Alta | | ● | ● | | |
| Baja | | | | | ⊝ |
| Consistency | 0.92 | 0.94 | 0.91 | 0.92 | 0.95 |
| Raw coverage | 0.13 | 0.10 | 0.14 | 0.16 | 0.11 |
| Unique coverage | 0.07 | 0.02 | 0.04 | 0.06 | 0.03 |

Note: ● = core causal condition (present); ⊖ = core causal condition (absent); • = contributing causal condition (present); ⊝ = contributing causal condition (absent).

Fuente: Elaboración propia.

Esta tabla, también informa de cálculos de *consistencia (Consistency)*, *cobertura bruta* (*Raw coverage*) y de *cobertura única* (*Unique coverage*), para cada una de las cinco soluciones.

**40.** Los resultados también se pueden resumir, con la ayuda de una tabla que clasifica las diferentes soluciones para evitar la disminución de *capacidad creativa de la empresa* **(A),** según el estado al que se dedica la muestra de empresas innovadoras clúster. (Ver **Tabla 9.6**).

**Tabla 9.6. Resultados del análisis del conjunto de datos difusos (*fuzzy-sets*) para lograr evitar la disminución de capacidad creativa de la empresa A, de acuerdo al estatus de la empresa**

| Estatus de la empresa | Evitar la disminución de capacidad creativa (A) |
|---|---|
| (P)=1 , (~B)=0 | (C) |
| (~P)=0 , (~B)=0 | (~S) * ((~C) + (C) * (I)) |
| (P)=1 , (B)=1 | (~S) * ((~C) ) * (~I) + (C) * (I)) |
| (~P)=0 , (B)=1 | {Ø} |

Fuente: Elaboración propia.

**41.** La **Tabla 12.6**, muestra que las diferentes soluciones son claramente evidentes para los CEOs provenientes de universidades públicas en diferentes categorías del estatus de las empresas innovadoras clúster. Aquellos casos de empresas que están en **P** y **(~B)** tienen más facilidad para evitar la disminución *de capacidad creativa de la empresa,* **A**. Todo lo que se requiere es una *educación de pregrado,* **C.**

**42.** En el otro extremo, no hay trayectorias causales o soluciones para evitar la disminución de *capacidad creativa de la empresa,* **A** para los CEOs quienes han cursado su educación en universidades públicas pero la empresa no se dedica a *alta tecnología* **(~P),** haciendo *productos consi-derando el impacto social,* **B.**

**43.** Igualmente, para evitar la disminución de capacidad creativa de la empresa **A,** para los CEOs quienes han cursado su educación en universidades públicas pero la empresa *no se dedica a alta tecnología* **(~P)** y *sin hacer productos considerando el impacto social* **(~B),** la solución es combinar puntajes de *capacidad innovadora* **S** no bajos, con *educación del CEO* **C** (pregrado o posgrado) combinada con altos *ingresos por innovación,* **I.**

**44.** Por otro lado, para evitar la disminución de *capacidad creativa de la empresa* **A**, para los CEOs quienes han cursado su educación en univer-sidades públicas pero la empresa *se dedica a alta tecnología* **P** y hacer *productos considerando el impacto social* **B**, la solución es similar, pero un poco más compleja, se deben combinar puntajes no bajos de *capa-cidad innovadora* **S** con ya sea:

  **a.** Valor de la *educación del CEO* **C,** de posgrado y con no bajos niveles de *ingresos por innovación,* **I** o,

  **b.** Valor de la *educación del CEO* **C,** de pregrado y con altos niveles de *ingresos por innovación,* **I**.

Como se observa, la clave es no tener baja *capacidad innovadora* (**S**), lo que indica a su vez que, aunque fuera pequeña o modesta, sigue siendo adecuada en el mundo de hoy.

## Finalizando

Los resultados presentados en el caso hipotético, son hallazgos preliminares, cuyo objetivo principal, es proporcionar al lector argumentos de contraste entre los análisis de *efectos netos* y los de tipo *configuracional*, con los mismos datos. El contraste entre los dos enfoques, debe ser claro. Los hallazgos de los análisis de *efectos netos* se expresan en términos de variables separadas. Proporcionan la cuenta final de contribuciones de los factores, para explicar la variación en el resultado, evitar la disminución de *capacidad creativa de la empresa,* **A**. La *educación de CEO* **C,** (de

pregrado o posgrado) y la dedicación de la empresa en la *alta tecnología* **P,** ganan esta contribución factorial, pero no elimina la *capacidad innovadora* **S,** ya que retiene un modesto efecto neto, a pesar de la dura interacción entre los factores (compárese con la **Tabla 9.1** y la **Tabla 9.2**). Los resultados de la regresión logística no mencionan el tema de las combinaciones causales, cuyo análisis requeriría el examen de modelos de interacción complejos. Examinando un modelo de interacción saturado, por ejemplo, requeriría la estimación de treinta y dos coeficientes en una sola ecuación. Incluso si tal modelo pudiera ser estimado (la colinealidad extrema hace que esta tarea sea inviable), el modelo sería prácticamente imposible de interpretar, una vez estimado. Tenga en cuenta también, que los supuestos de aditividad y linealidad en el análisis de regresión logística, permite la estimación de probabilidades de resultado para los treinta y dos sectores del espacio vectorial, definidos por las cinco variables independientes. Por lo tanto, el enfoque de *efectos netos*, aborda el problema de la diversidad limitada de manera indirecta y encubierta, asumiendo que el efecto de una variable dada es el mismo, independientemente de los valores de las otras variables y que una relación lineal, se puede extrapolar más allá de un rango de valores observado. Para generar la probabilidad estimada de evitar la disminución de *capacidad creativa de la empresa* **A**, en cualquier punto del espacio vectorial definido por las variables independientes, es necesario, simplemente insertar las coordenadas de ese punto en la ecuación y calcular el valor predicho. La cuestión de la diversidad limitada, se deja de lado por completo. Por el contrario, este problema debe confrontarse en un análisis de configuración. Los datos que ocurren naturalmente, están profundamente limitados en su diversidad como se ilustra en la **Tabla 9.3**. Este hecho es evidente, cada vez que los investigadores examinan la distribución de casos, a través de las combinaciones causales lógicas posibles, especialmente, cuando el número de las *condiciones causales*, son más que unas pocas. Como lo ilustra el análisis aquí presentado, el problema de la diversidad limitada, no se soluciona concibiendo una gran cantidad de casos.

Cuando los casos se visualizan de forma *configuracional*, es posible identificar las diferentes combinaciones de *c ondiciones causales* vinculadas a un resultado. Los resultados del análisis *configuracional*, reportados en este capítulo, muestran que hay varias soluciones o trayectorias causales para evitar la disminución de *capacidad creativa de la empresa* **A**, para los CEOs con educación en universidades públicas de las empresas innovadoras clúster de México. Todas las soluciones o trayectorias causales, incluyen calificaciones en la educación del CEO **C** (*pregrado y posgrado*) y un estatus de empresa favorable (ya sea *dedicado o no a la tecnonlogía* **P**, o *haciendo no productos con impacto social,* **B**). No tener bajas las puntuaciones en *capacidad innovadora* **S,** también son una condición en cuatro de las cinco soluciones o trayectorias causales, como ingresos en *innovación altos o no bajos* **I** y en estos mismos, tres soluciones o trayectorias causales. Un investigador podría dramatizar las implicaciones, afirmando que si se pudiera elegir, entre tener un *puntaje de capacidad innovadora* **S,** alto y un *ingreso por innovacion alto* **I,** la mejor opción sería seleccionar un alto puntaje de *capacidad innovadora* **S.** Los resultados de *datos difusos*, subrayan el hecho de que la elección se trata realmente de combinaciones de condiciones, de soluciones o trayectorias causales, no de variables individuales. En resumen, elegir no tener un puntaje de *capacidad innovadora* **S** baja, por sí solo, no ofrece protección contra evitar la disminución de *capacidad creativa de la empresa* **A**, por lo que el análisis *con igguracional* presentado en este capítulo muestra claramente, que se combinan con otros recursos.

## Resumen fsQCA de calibraciones usadas en el análisis de conjunto de datos difusos (*fuzzy-sets*)

Como se señaló anteriormente, la calibración es esencial para todo análisis del conjunto de *datos difusos*. Las calibraciones incorrectas distorsionan los resultados de las evaluaciones del marco teórico de la investigación. Los principios principales que guían la calibración son:

**a.** El objetivo del conjunto de *datos difusos*, debe estar cuidadosamente bien definido y etiquetado, y

**b.** Las puntuaciones del conjunto de *datos difusos*, debe reflejar los estándares externos, basados tanto en la existencia del conocimiento empírico sustantivo del investigador, así como en el diseño del marco teórico de la investigación.

Mientras algunos podrían considerar que, la influencia de las decisiones de calibración indebidas y llevar este aspecto del análisis difuso, como una debilidad, de hecho, se considera como una fortaleza. Esto es, porque los investigadores deben prestar mucha atención a la definición y la construcción de sus conjuntos de *datos difusos*, y se ven obligados a conceder que el conocimiento sustantivo es, en esencia, un requisito previo para análisis por lo que la *calibración* es un paso importante. Los conjuntos de *datos difusos* en el análisis presentado en este capítulo, son de grado de pertenencia en el resultado **A**, así como de grado de pertenencia, en conjuntos que reflejan diversos antecedentes, características y condiciones, tales como:

**1.** *Evitar la disminución de capacidad creativa de la empresa* **A.** Para construir este conjunto difuso la cual, es considerada la variable dependiente binaria en el análisis de regresión logística. Sin embargo, por su naturaleza dicotómica, coloca a las empresas del clúster *con ingresos por innovación* **I,** apenas superiores, en la misma categoría de aquellas empresas con ingresos notablemente muy por encima del umbral de *ingresos por innovación* **I.** El procedimiento del conjunto de *datos difusos* al usar el método directo para calibrar el conjunto de *datos difusos* (ver **Tabla 9.1**), muestra el umbral para la pertenencia completa en el conjunto de empresas que no están en **A** (puntaje difuso = **0.95**) en una proporción de **3.0** (el *ingreso de innovación* **I** es tres veces el nivel de, **A**), el punto de cruce (puntaje difuso = **0.5**) es una relación de **2.0** (el *ingreso de innovación* **I** es el doble del nivel de **A**) y el umbral para la exclusión total del conjunto de empresas que no están **A** (puntaje difuso= **0.05**) es una razón de **1.0** (el *ingreso de innovación*, **I** es el mismo que el nivel de **A**).

2. *Educación del CEO (pregrado y posgrado)* **C**. Para medir el logro educativo, esta variable traduce los años de educación, directamente en grados (es decir, completar dieciseis años de educación indica un título de pregrado, mientras veintidos años completados, indica un título de posgrado). Los encuestados con trece o más años de escuela están completamente en el conjunto pregrado (una puntuación difusa de **1.0**). Por otro lado, aquellos con solo educación de bachillerato (es decir, doce años de escuela o menos) se tratan como totalmente fuera del conjunto de los de pregrado (un puntaje difuso de **0.0**). El conjunto difuso abarca los 15 años= **0.75**; 14 años = **0.60**; 13 años = **0.45**; 12 años = **0.30**; y 11 años = **0.15**. El conjunto de datos difuso de los encuestados con educación de posgrado, se construyó definiendo encuestados con veintidos o más años de educación que tienen pertenencia completa en el set con educación de posgrado **(1.0),** mientras que aquellos con dieciocho años de educación o menos fueron codificados como totalmente fuera del conjunto **(0.0).** Los años intermedios se codificaron de la siguiente manera: 21 años= **0.60**; 20 años= **0.40** y 19 años = **0.20.**
3. *Ingresos por innovación***, I.** Son datos hipotéticos, producidos como el promedio de ingreso (en mdp) que reportan las empresas innovadoras del clúster, a partir del año cuatro, de acuerdo a los recursos invertidos para innovar en distintos rubros como: modelo de negocios, productos/servicios, mercadotecnia, procesos, organización **(OECD, 2018; Mejía-Trejo et al., 2013)**. Estos datos fueron utilizados para crear dos conjuntos de *datos difusos*: el conjunto de encuestados con bajos ingresos y el de altos *ingresos por innovación,* **I**. Usando el *método directo de calibración* descrito, en la **Tabla 8.1,** se determinó que el umbral para la pertenencia completa en el conjunto con bajos ingresos de los padres **(0.95)** es una relación de **1.0** (Evitar la disminución de *capacidad creativa de la empresa,* **A**). Encuestados con ratios menores de **1.0** recibió puntajes difusos mayores de **0.95**. Por el contrario, el umbral para la exclusión total **(0.05)** del conjunto con bajos ingresos parentales es una proporción de **3.0** (el *ingreso por innovación* **I**, era tres veces el nivel, **A**). Los encuestados con relaciones superiores a **3.0** reci-

recibieron puntajes difusos menores a **0.05.** Se determinó que el punto de cruce era dos veces el nivel de **A.** Múltiplos de la razón de **A** (*ingreso de innovación* **I)** dividido por nivel **A** ajustado por la composición de empresas) también se utilizaron para construir el conjunto difuso de encuestados con altos *ingresos de innovación* **I.** El umbral de exclusión del conjunto de empresas con altos *ingresos de innovación* **I (0.05)** es una proporción de tres veces el nivel de **A** ajustado. El *punto de cruce* **(0.50)** se estableció en **5.5** veces el nivel de **A** ajustado y el umbral para la pertenencia completa se estableció en ocho veces el nivel ajustado de **A**. El *umbral para la pertenencia completa* corresponde aproximadamente a *tres veces* el *ingreso por innovación* **I** de empresas promedio, mientras que el punto de cruce corresponde aproximadamente dos veces el *ingreso por innovación* **I** de empresas promedio. De nuevo, el método directo de calibración del conjunto de *datos difusos*, se utilizó para calibrar el grado de pertenencia.

4. *Capacidad innovadora* **S**. Las puntuaciones son propuestas hipotéticas de acuerdo al modelo sugerido en **Mejía-Trejo (et al., 2013)**, basado en cinco categorías. Estas cinco categorías, tienen una importancia sustancial y determinan la elegibilidad y la asignación a diferentes calificaciones grupos de empresas que basan su capacidad innovadora en el Manual de Oslo **(OECD, 2018)**, en rubros como: modelo de negocios, productos/servicios, mercadotecnia, procesos, y organización. Así, la **categorías I** (percentil **93** al **99**) y **II** (percentil **65** a **92**) se consideran superiores a la media en capacidad de entrenamiento; los de la **categoría III** (percentil **31** al **64**) son aproximadamente promedio; los de la **categoría IV** (percentil **10** al **30**) se designan como inferior a la media en capacidad de entrenamiento; y los de la **categoría V** (percentil **1** al **9**) están marcadamente por debajo del promedio. Además, los que puntúan en la **categoría IV** (percentil **10** al **30**) no son elegibles para alistarse a menos que el CEO tenga una educación de pregrado. Además se preve que no más del 20% de los enlistados se extraen de la **categoría IV**, lo que indica además que los encuestados en esta categoría son sustan-

cialmente diferentes de aquellos en **categorías I** a **III.** Para construir el conjunto de datos difuso de encuestados con puntajes **S** bajos, se utilizan los puntajes percentiles **S** de los encuestados. El umbral para la plena pertenencia **(0.95)** en el conjunto de encuestados con puntajes (**S**) bajos fue colocado en el **percentil 10**; los encuestados que obtuvieron puntajes inferiores al **percentil 10** recibieron puntuaciones de *datos difusos* de pertenencia mayores a **(0.95)**. Se estableció el punto de cruce (**0.5**) en el **percentil 20**, y el umbral para la no pertenencia se estableció en el **percentil 30**. Los encuestados que obtuvieron mejores puntajes que el **percentil 30** recibió puntajes difusos de menos de **0.05** en el conjunto de encuestados con bajo puntaje **S.** El umbral para la pertenencia completa **(0.95)** en el conjunto de encuestados con puntuaciones **S** altos se colocó en el **percentil 93**, en línea con la designación militar del límite inferior del más alto categoría; el punto de cruce **(0.5)** se estableció en el **percentil 80**; y el umbral para la no pertenencia completa **(0.05)** en el conjunto de encuestados con puntajes **S** altos se colocó en el **percentil 65**.

**5.** *Composición de las empresas*. Tiene dos componentes principales: si la empresa se dedica *no a la alta tecnología* **P** y *si sus productos son o no de impacto social* **B.** Las cuatro combinaciones de alta tecnología / no alta tecnología e impacto social / no impacto social, están presentes con frecuencia en el conjunto de datos. El caso de dedicación a la alta tecnología de los encuestados, está codificado como datos nítidos, con un valor de uno asignado a aquellos que fueron creados en 1990. En general, las empresas dedicadas a la alta tecnología, son mucho menos propensas estar en la *pérdida de creatividad* **A.** Tener *productos con impacto social* **B** se codifica aquí como conjunto de datos nítido. La razón es que considerar el impacto social, impone cierto estatus y restricciones en el actuar de la empresa, en el que el primer producto es más importante que los sucesivos en la percepción del público consumidor en general. Así, las empresas que fabrican productos de *impacto social* **B,** son más probables que busquen realizar actividades que eviten la *pérdida de*

*creatividad* **A** que las empresas que no los fabrican. La composición de empresa, con respecto a mantenerse fuera de **A,** es la combinación de dedicación a alta tecnología **P** / *sin productos de impacto social*, **B.** El menos favorable, es la combinación *no dedicación a alta tecnología* **P** / *productos de impacto social*, **B.**

## CAPÍTULO 10.
# Guía de análisis csQCA

El objetivo de esta sección es proveer una guía práctica *paso a paso* sobre el uso del software **fsQCA 3.0** para análisis **csQCA** aplicando los pasos descritos anteriormente. Así, tomaremos de referencia el trabajo de **Mejía-Trejo y Rodríguez-Bravo (2019)** titulado: *Values-Based Innovation. Designing a model to be Applied in Management Sciences.* Este modelo, es base de propuesta para detectar las empresas que practican innovación abierta **OIN** en México, basados en el modelo de innovación por valores, relacionando a los siguientes factores:

**a.** Innovación normativa (etapa de producción que incorpora el valor de la empresa y regulaciones a la que se somete como propósito general) **NIN.**

**b.** Innovación en el modelo de negocios (etapa de consumo, que incorpora e valor corportativo, los propósitos estratégicos de la empresa), **IMN.**

**c.** Innovación en el producto/servicio diferenciado (etapa de diferenciación, que incorpora los valores de la organización, individualidad que se proyectan en el producto/servicio a mostrar a sus consumidores), **DIF.**

**d.** El modelo de evaluación de valores (califica los valores anteriores), **EVA.**

Del que se deriva, la siguiente pregunta de investigación: *¿por qué algunos empresas han experimentado innovación abierta inspirada en los valores de la misma, mientras otros no?*

Se tienen **39** casos empresas que son consideradas practicantes de innovación abierta en México cuyo resultado de interés es: *determinar las empresas con efectividad en la práctica de la innovación abierta* **OIN**, *a través del modelo de valores de innovación.* Ver **Tabla 10.1.**

**Tabla 10.1. Matriz hipotética de datos dicotómicos a partir del estudio de Mejía-Trejo y Rodríguez-Bravo (2019)**

| Empresas que practican innovación abierta Casos | Condiciones | | | | Resultado |
|---|---|---|---|---|---|
| | NIN | IMN | DIF | EVA | OIN |
| 1 | 1 | 1 | 0 | 0 | 1 |
| 2 | 1 | 1 | 0 | 0 | 1 |
| 3 | 1 | 1 | 0 | 0 | 1 |
| 4 | 1 | 1 | 0 | 0 | 1 |
| 5 | 1 | 1 | 0 | 0 | 1 |
| 6 | 1 | 1 | 0 | 0 | 1 |
| 7 | 0 | 1 | 0 | 0 | 1 |
| 8 | 0 | 1 | 0 | 0 | 1 |
| 9 | 0 | 1 | 0 | 0 | 1 |
| 10 | 1 | 1 | 0 | 1 | 1 |
| 11 | 1 | 0 | 0 | 0 | 1 |
| 12 | 1 | 1 | 1 | 0 | 0 |
| 13 | 1 | 1 | 1 | 0 | 0 |
| 14 | 1 | 1 | 1 | 0 | 0 |
| 15 | 1 | 1 | 1 | 0 | 0 |
| 16 | 1 | 0 | 1 | 0 | 0 |
| 17 | 1 | 0 | 1 | 0 | 0 |
| 18 | 1 | 0 | 1 | 0 | 0 |
| 19 | 1 | 0 | 1 | 0 | 0 |
| 20 | 1 | 0 | 1 | 0 | 0 |
| 21 | 1 | 0 | 1 | 0 | 0 |
| 22 | 0 | 0 | 0 | 0 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 |
| 24 | 0 | 0 | 0 | 0 | 0 |
| 25 | 0 | 1 | 0 | 1 | 0 |
| 26 | 0 | 1 | 0 | 1 | 0 |
| 27 | 0 | 1 | 0 | 1 | 0 |
| 28 | 0 | 1 | 0 | 1 | 0 |
| 29 | 0 | 1 | 0 | 1 | 0 |
| 30 | 0 | 0 | 0 | 1 | 0 |
| 31 | 0 | 0 | 0 | 1 | 0 |
| 32 | 0 | 0 | 0 | 1 | 0 |
| 33 | 0 | 0 | 0 | 1 | 0 |
| 34 | 0 | 0 | 0 | 1 | 0 |
| 35 | 1 | 0 | 1 | 1 | 0 |
| 36 | 1 | 0 | 1 | 1 | 0 |
| 37 | 0 | 1 | 1 | 1 | 0 |
| 38 | 0 | 0 | 1 | 1 | 0 |

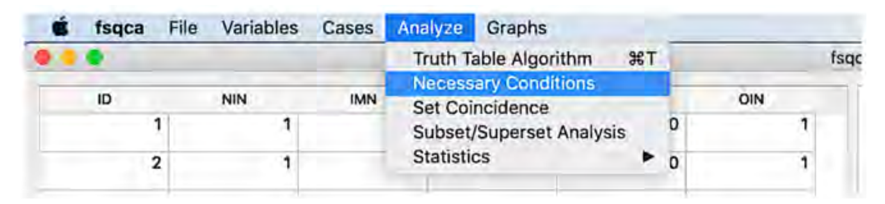
Fuente: Elaboración propia.

A partir de este diseño se va a proceder a reproducir el análisis de **QCA** paso a paso con el software **fsQCA 3.0**

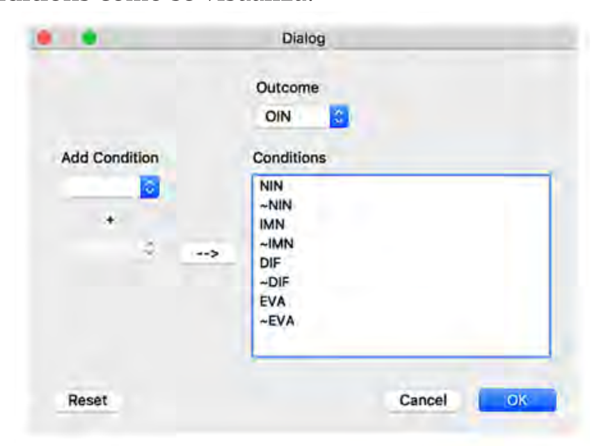
## Análisis de condiciones necesarias. Caso 1

Para las pruebas de necesidad, se va a comenzar con el resultado **OIN**, por lo que del archivo de trabajo: **value based innovation.csv**

1. Oprima: **Analyze → Necessary conditions** para, en la casilla **Outcome**, seleccionar como resultado **OIN**.




2. Seguidamente, en la pestaña **Add Condition**, se seleccionan las condiciones (y sus complementos) que se incorporan al contenedor referente a las **conditions** como se visualiza:



3. Una vez seleccionado todo, oprimir **OK**, que genera los resultados para cada condición, como se muestra:

```
Analysis of Necessary Conditions

Outcome variable: OIN

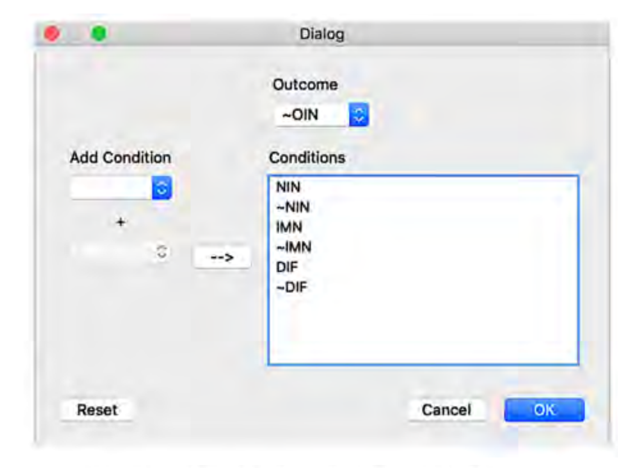
Conditions tested:
       Consistency     Coverage
NIN    0.727273        0.400000
~NIN   0.272727        0.166667
IMN    0.909091        0.500000
~IMN   0.090909        0.055556
DIF    0.000000        0.000000
~DIF   1.000000        0.458333
EVA    0.090909        0.066667
~EVA   0.909091        0.434783
```

4. Según lo expuesto, una condición puede ser considerada como *necesaria* si cumple el criterio de necesidad por encima del *umbral de consistencia de* **0.9**.
5. Como se observa, hay tres condiciones que pasan este umbral de consistencia **(IMN, ~ DIF y ~ EVA).** Se ha aplicado un umbral algo más conservador al nivel de **(1.0)** para seleccionar las *condiciones necesarias* en el estudio, considerando solo **(~ DIF)** como *necesaria* para el resultado.
6. Esta expresión *indica necesidad, que no causalidad*, y se lee del siguiente modo: la ausencia de **(~ DIF)** (*innovación en el producto/servicio diferenciado*) es una condición necesaria para la ocurrencia la *efectividad de la innovación abierta,* **OIN**.
7. Esta información es *relevante* para el análisis de *suficiencia* que se ejecutará en los próximos apartados, puesto que dicha condición no debe ser incluida en la *minimización* para el resultado **(~ OIN)** como parte de un análisis estándar mejorado.
8. Así, considerando el valor **(1.0)** de *consistencia*, se obtiene:

$$(\sim\mathbf{DIF}) \rightarrow \mathbf{OIN}$$

9. Un segundo paso en el *análisis de necesidad* es su aplicación a la *no ocurrencia del fenómeno*. El procedimiento es exactamente el mismo

que el anterior, con la diferencia de que en este caso se incluye como resultado **(~ OIN).** Los datos arrojados por el software son:




```
Analysis of Necessary Conditions

Outcome variable: ~OIN

Conditions tested:
        Consistency     Coverage
NIN     0.444444        0.600000
~NIN    0.555556        0.833333
IMN     0.370370        0.500000
~IMN    0.629630        0.944444
DIF     0.518519        1.000000
~DIF    0.481481        0.541667
```

**10.** De acuerdo a los datos entregados por el software, *no hay condiciones necesarias* para la ausencia del resultado, porque *ninguna condición pasa el umbral* de **(1.0)** escogido en el diseño por los autores.

**11.** Tal como se explicó en este capítulo, si las *pruebas de necesidad* no arrojan resultados positivos para las *condiciones individuales*, una estrategia es *formar equivalentes funcionales* y aplicar el *análisis de necesidad a cada uno de ellos*. Sin embargo, no se aplicó esta estrategia en el estudio.

**12.** Para efectos de ilustración, se realizará un *análisis de necesidad* de *equivalentes funcionales* de cada posible unión de condiciones, las cuales mostramos:




Analysis of Necessary Conditions

Outcome variable: ~OIN

Conditions tested:

| | Consistency | Coverage |
|---|---|---|
| NIN+IMN | 0.666667 | 0.620690 |
| NIN+~IMN | 0.777778 | 0.724138 |
| ~NIN+DIF | 1.000000 | 0.900000 |
| NIN+~DIF | 0.925926 | 0.694444 |
| NIN+EVA | 0.888889 | 0.750000 |
| NIN+~EVA | 0.555556 | 0.576923 |
| IMN+DIF | 0.703704 | 0.655172 |
| IMN+~DIF | 0.666667 | 0.620690 |
| IMN+EVA | 0.666667 | 0.642857 |
| IMN+~EVA | 0.703704 | 0.633333 |
| ~IMN+DIF | 0.814815 | 0.956522 |
| ~IMN+~DIF | 0.814815 | 0.666667 |
| ~IMN+EVA | 0.851852 | 0.920000 |
| ~IMN+~EVA | 0.777778 | 0.677419 |
| DIF+EVA | 0.888889 | 0.960000 |
| DIF+~EVA | 0.629630 | 0.629630 |
| ~DIF+EVA | 0.629630 | 0.607143 |
| ~DIF+~EVA | 0.851852 | 0.676471 |

**13.** De acuerdo a los resultados, la unión de **(~ NIN + DIF)** sería una condición necesaria para **(~OIN)** teniendo una *consistencia de* **(1.0)** y *cubriendo* el **90%** de los casos.

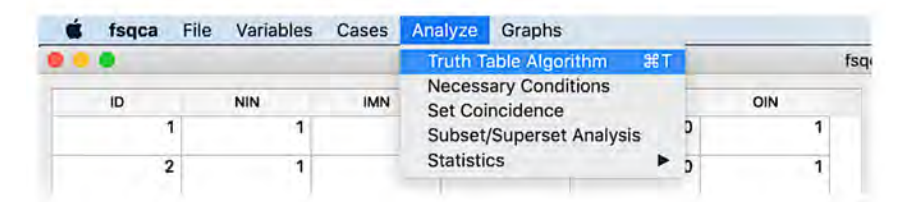
14. Otra unión con consistencia por encima del umbral de **(0.9)** es **(NIN + ~ DIF),** cubriendo el **70%** de los casos.
15. No obstante, esta última *contradice* a la primera unión, por lo que para efectos de una posible selección del *equivalente funcional*, se podría considerar como *necesario* únicamente a aquel que contemple la *consistencia y cobertura más relevantes*.
16. Es importante insistir, una vez más, que para seleccionar uniones como condiciones necesarias debe existir una argumentada justificación teórica y empírica.
17. En cualquier caso, los resultados de nuestro análisis de necesidad, a modo de ejemplo, para **(~ OIN)** indican que la unión de la *ausencia* de **NIN (~ NIN)** con la *presencia* de **DIF** es una condición necesaria para la *ausencia* de **(~ OIN)**, que se representaría del siguiente modo:

**(~NIN) + (DIF) → (~ OIN)**

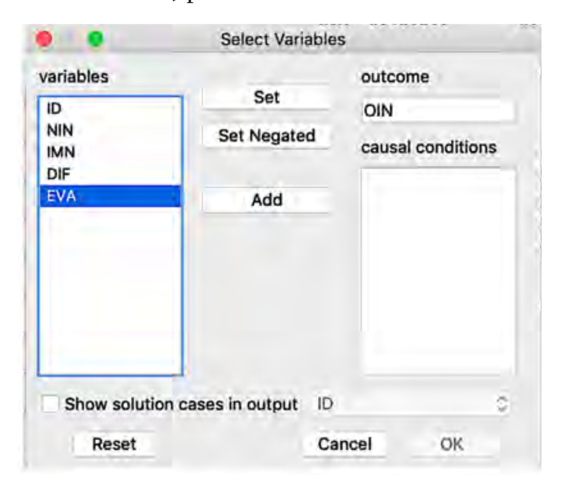
## Representar datos en la tabla de verdad. Caso 2

Para transformar el conjunto de datos de la matriz dicotómica en una *tabla de verdad*, se debe:

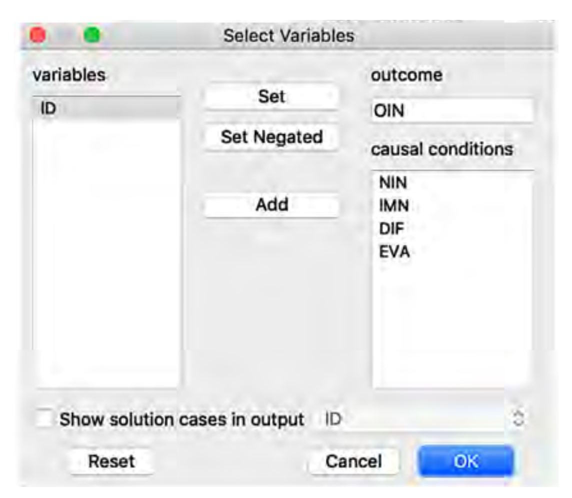
1. Oprimir en: **Analyze → Truth Table Algorithm**, generando una ventana como la mostrada:



2. En la nueva ventana, se debe especificar el resultado **OIN** y desplazar hacia la casilla **Outcome**, pulsando el botón **Set**.

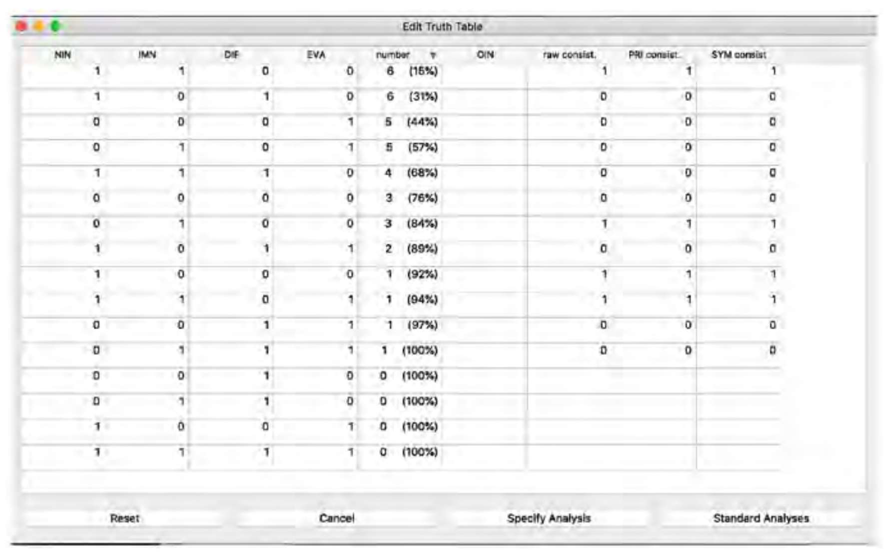



3. Para las condiciones, se deben seleccionar aquellas de interés contenidas en la casilla **Variables** y desplazar hacia **Causal Condition** oprimiendo en el botón **Add**.



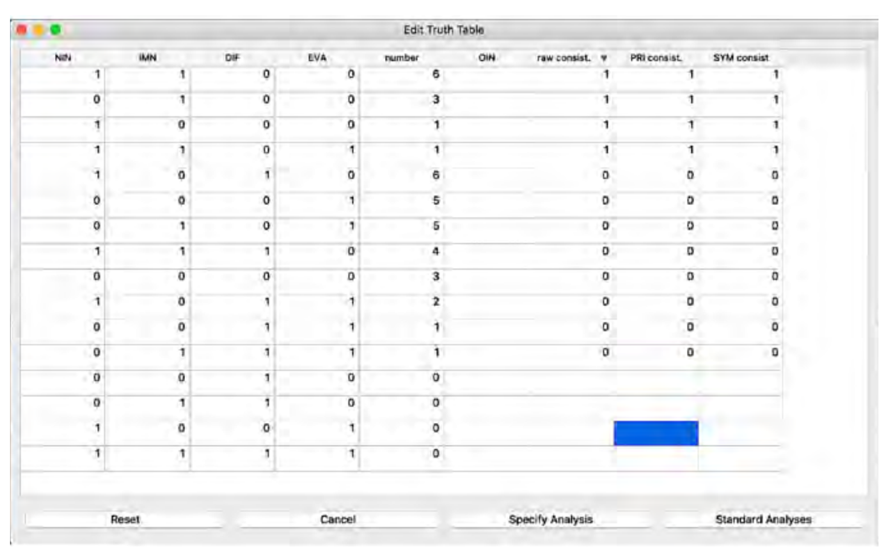

4. Después de pulsar el botón **OK** la *tabla de verdad* aparece en una nueva ventana. Como puede apreciarse, esta consiste en $\mathbf{2}^{k}$ filas, siendo $\boldsymbol{k}$ el número de condiciones. Como son **cuatro las condiciones**, se producen

**16 filas**. La columna **number** indica el número de casos que tienen pertenencia mayor a **0.5** en una determinada fila. Para ordenar las filas de acuerdo a **raw consistency** (*consistencia bruta*), se debe oprimir en **Sort** y escoger la opción **descending** (*decreciente*).


Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 (15%) | | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 (31%) | | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 (44%) | | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 (57%) | | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 (68%) | | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 (76%) | | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 3 (84%) | | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 2 (89%) | | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 (92%) | | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 (94%) | | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 (97%) | | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 (100%) | | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 (100%) | | | | |
| 0 | 1 | 1 | 0 | 0 (100%) | | | | |
| 1 | 0 | 0 | 1 | 0 (100%) | | | | |
| 1 | 1 | 1 | 1 | 0 (100%) | | | | |

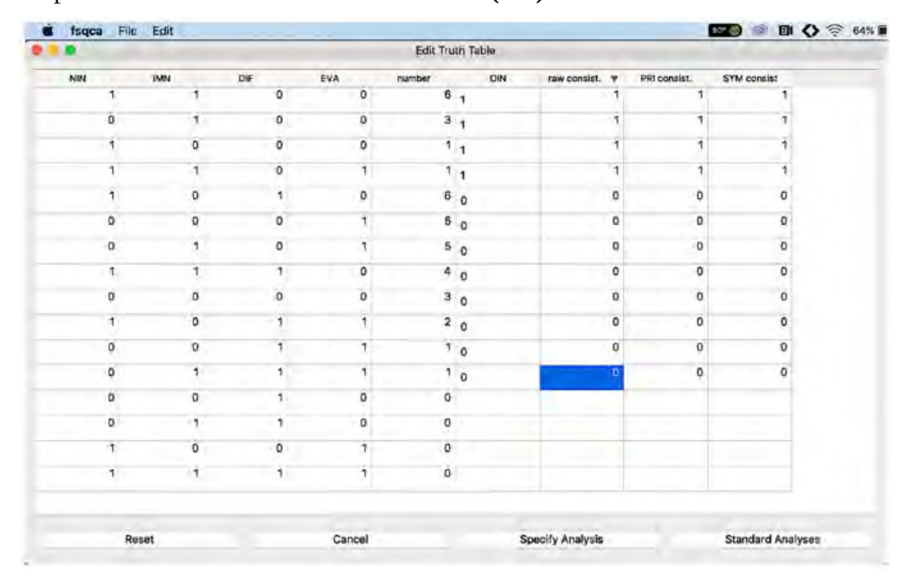
Reset | Cancel | Specify Analysis | Standard Analyses


Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 3 | | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 | | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 2 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | | | | |
| 0 | 1 | 1 | 0 | 0 | | | | |
| 1 | 0 | 0 | 1 | 0 | | | | |
| 1 | 1 | 1 | 1 | 0 | | | | |

Reset | Cancel | Specify Analysis | Standard Analyses

5. **5.** Como puede observarse, la columna del resultado **OIN** está vacía, porque antes se debe analizar la *suficiencia* de cada fila para el resultado.
6. **6.** Las filas que contienen bastantes números, que son subconjuntos del resultado y que tienen una **raw consistency** (*consistencia bruta*) de **(1.0)**, son suficientes para el resultado, tales como las primeras cuatro filas de la *tabla de verdad*.
7. **7.** Por lo tanto, a estas cuatro filas se les debe agregar el valor de **(1)** en la columna **OIN** y el valor de **(0)** en la misma columna a aquellas filas que presenten una consistencia inferior a **(1.0)**.

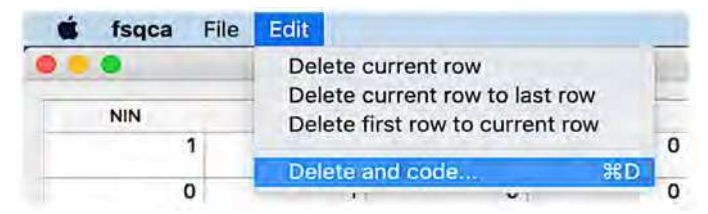


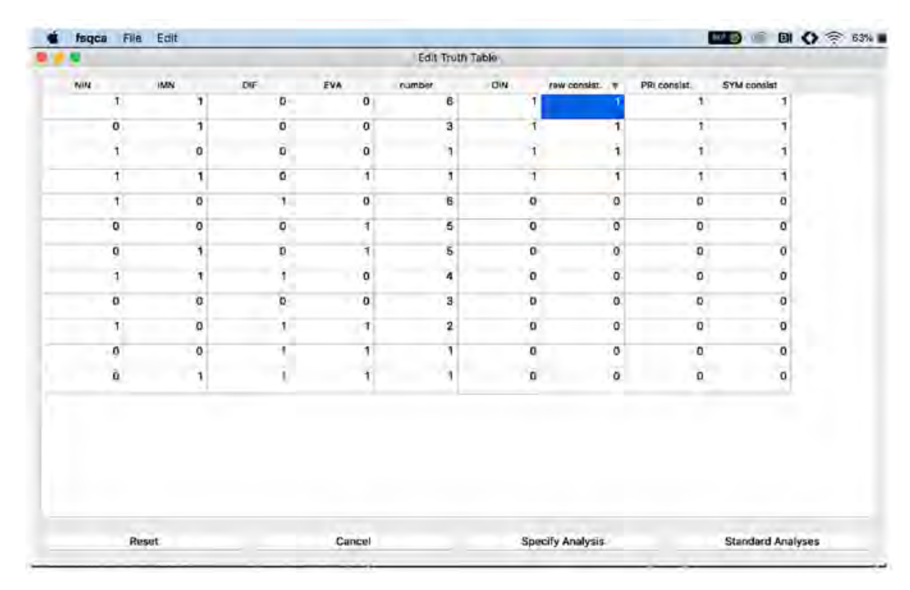

| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 2 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | | | | |
| 0 | 1 | 1 | 0 | 0 | | | | |
| 1 | 0 | 0 | 1 | 0 | | | | |
| 1 | 1 | 1 | 1 | 0 | | | | |

8. **8.** El software, por defecto, tiene un *umbral de suficiencia* de **0.8**, que puede ser modificado bajo justificación teórica y empírica por el investigador.



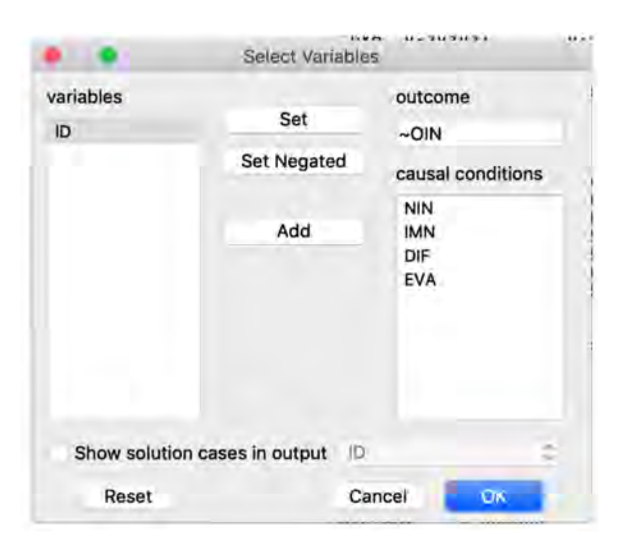
9. Al ir a la pestaña: **Edit → Edit and code...** aparecerá la siguiente ventana para *ajustar los umbrales de consistencia*.



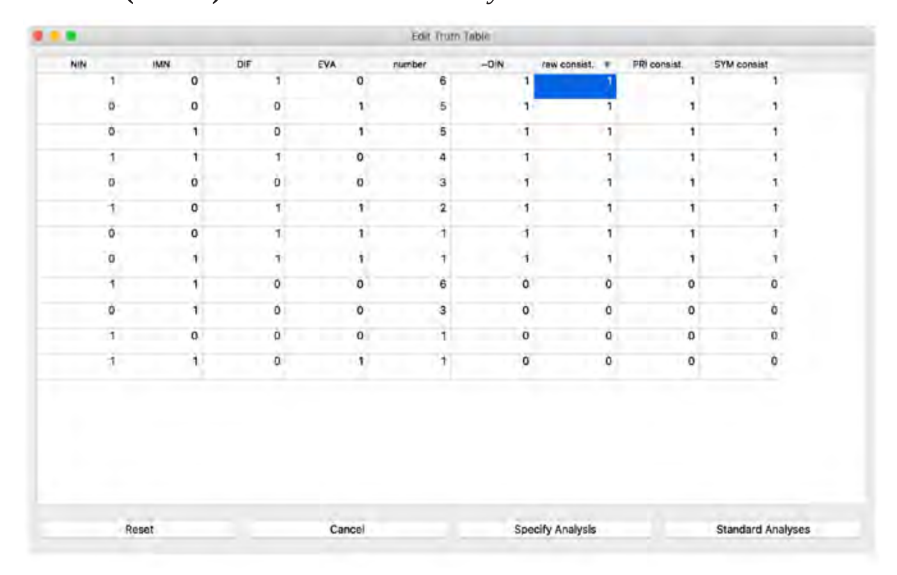




| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 2 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |

10. El procedimiento para la *no-ocurrencia* del resultado **(~ OIN)** es el mismo y la única diferencia es que al escoger el resultado desde la casilla *Variables* esta debe ser desplazada la casilla **Outcome** a través del botón **Set Negated**. Así, una nueva tabla de verdad ha sido creada con el conjunto de **(~ OIN).**




**11.** En la imagen se puede observar que existen ocho configuraciones que cumplen el **criterio de suficiencia (< 0.8)** para contribuir al resultado de interés **(~ OIN)** conteniendo ***27** casos y una consistencia de **1.0**.*



Edit Truth Table

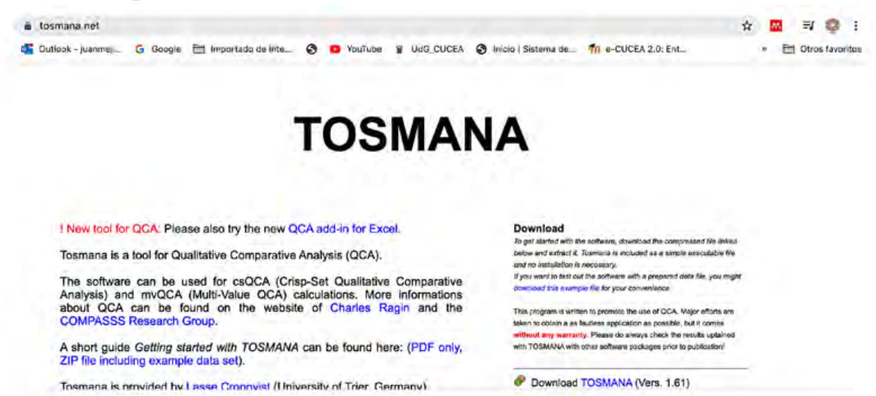
| NIN | IMN | DIF | EVA | number | ~OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 6 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 1 | 5 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 5 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 4 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |

Reset    Cancel    Specify Analysis    Standard Analyses

## Identificar filas contradictorias. Caso 3

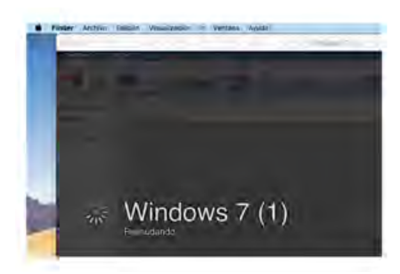
En la matriz de datos *dicotómicos*, no se observan contradicciones, es decir, no hay ninguna combinación de condiciones idénticas que contribuya tanto a la presencia del resultado como a su ausencia. Si bien en esta guía de **QCA** se introduce el análisis con el software **fsQCA**, hay un aspecto del software **TOSMANA**, que merece la pena presentar: la *visualización de contradicciones en un diagrama de Venn*, realice:

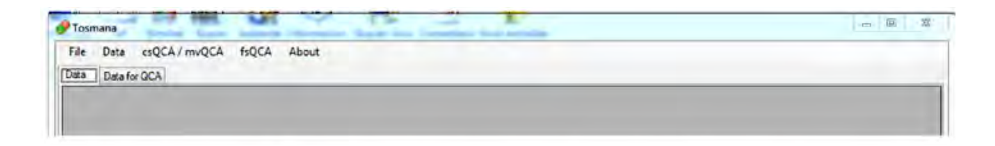
1. Localice y descargue en su máquina del sw TOSMANA, en el siguiente link: https://www.tosmana.net/



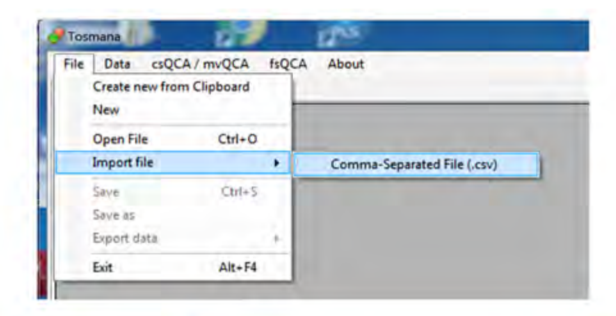

2. Descargue la versión **TOSMANA1.61**, el cual, al momento de escribir este documento, solo corre en sistema operativo Windows. Si usted maneja sistema operativo OS requerirá de un emulador como Parallels, para poder correrlo (como es nuestro caso).

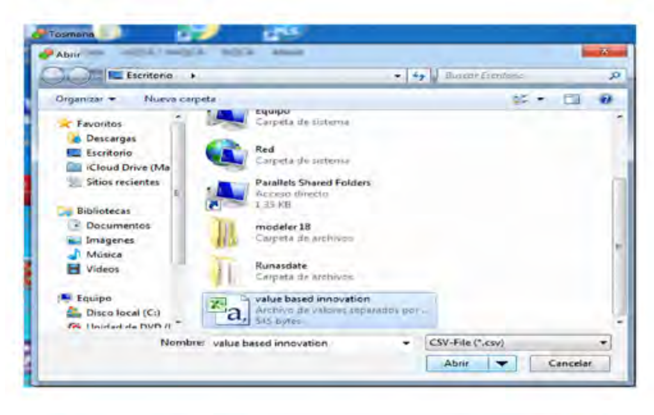


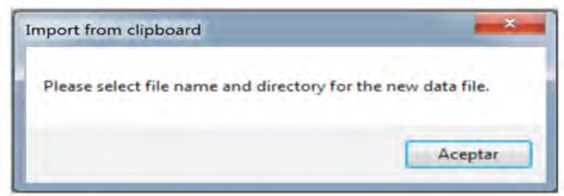



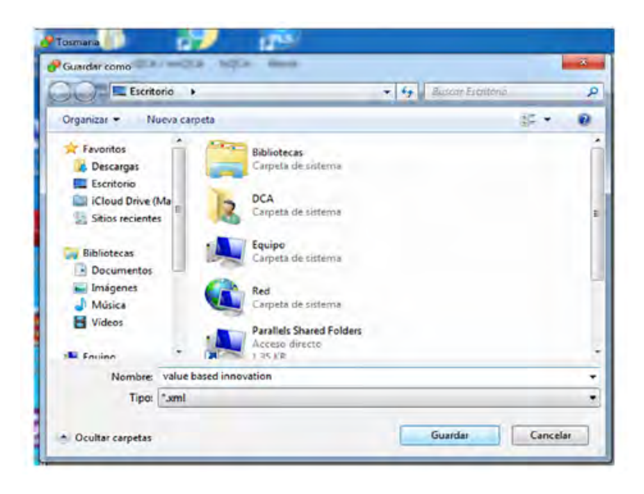



3. **TOSMANA 1.61** solo lee archivos en formato XML, por lo que nuestro archivo de trabajo: **value based innovation.cvs**, deberá importarse y realizar las actividades que sugerimos en las imágenes, para lograrlo. Para accesar e importar el archivo basta oprimir: **File → Import file → comma separated file (.csv)** y seleccionar el archivo o matriz de datos en cuestión.

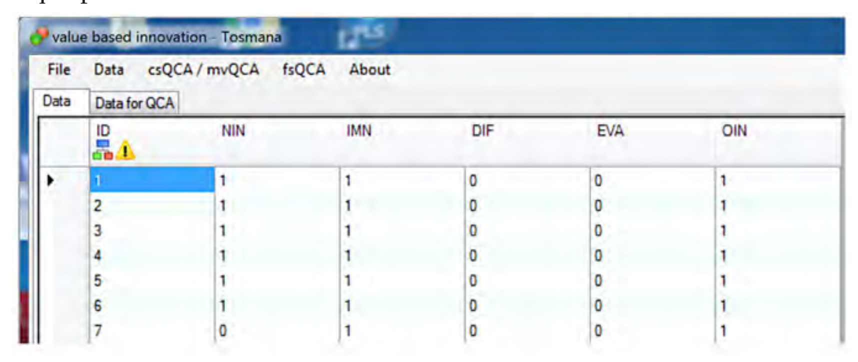


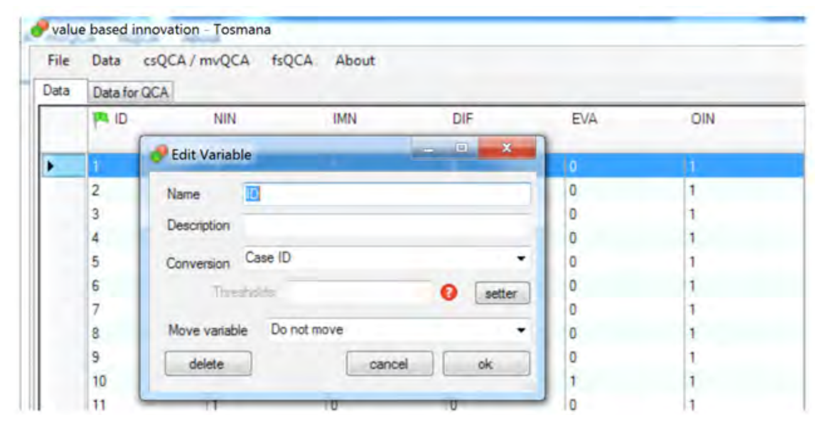



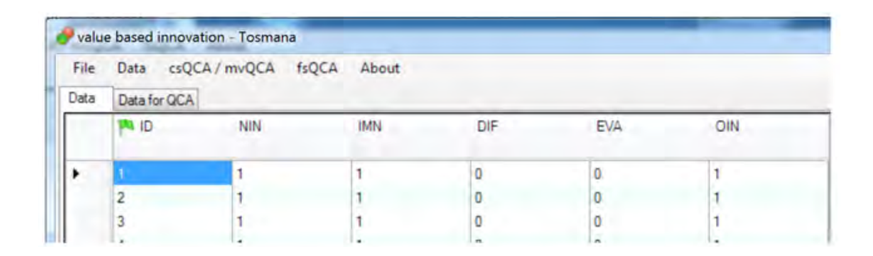





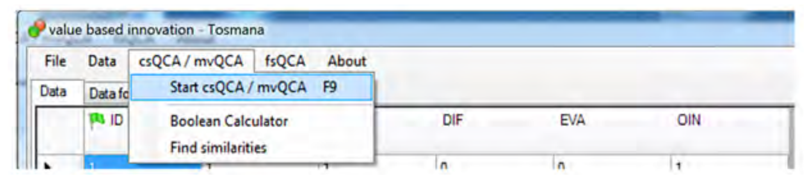

4. Nota: es muy importante declarar el ID dentro del archivo XML para que pueda referenciar los casos. Observe como cambia a bandera verde.

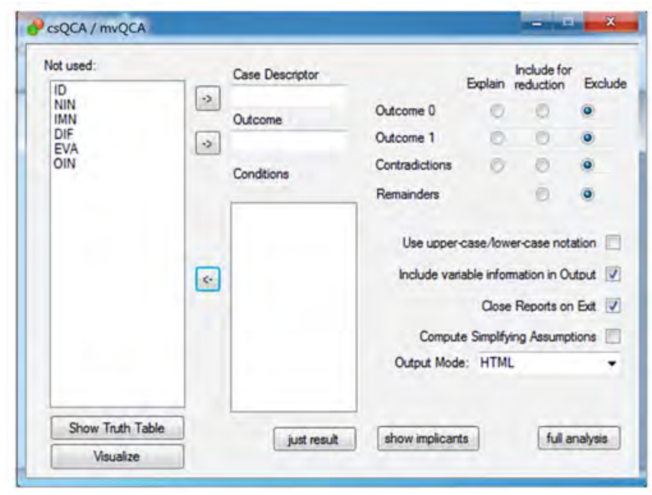


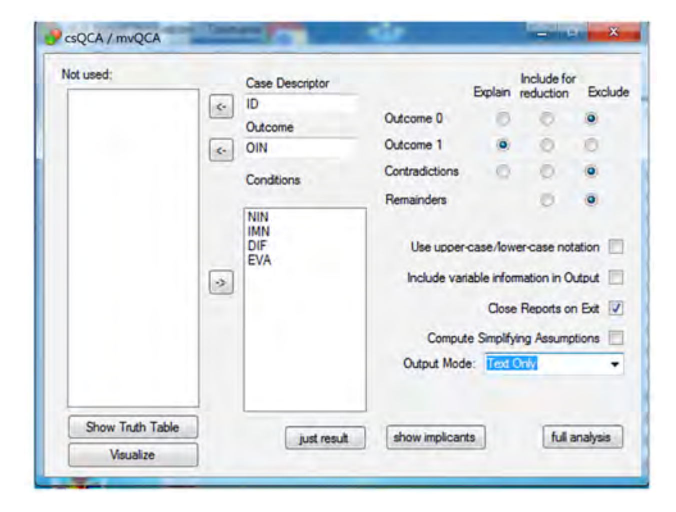



5. Oprima **Analysis → Start MV(QCA).**

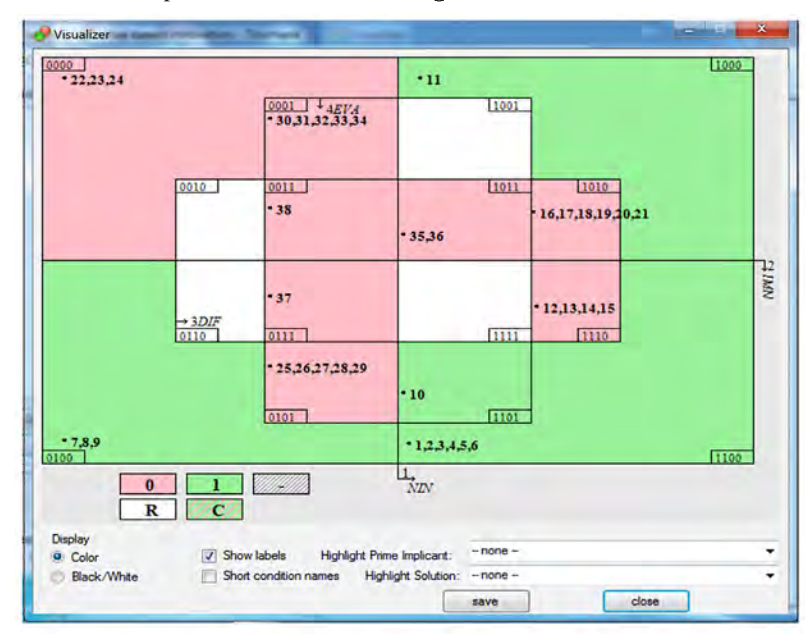


**6.** Hacer la selección de las casillas, como se muestra:

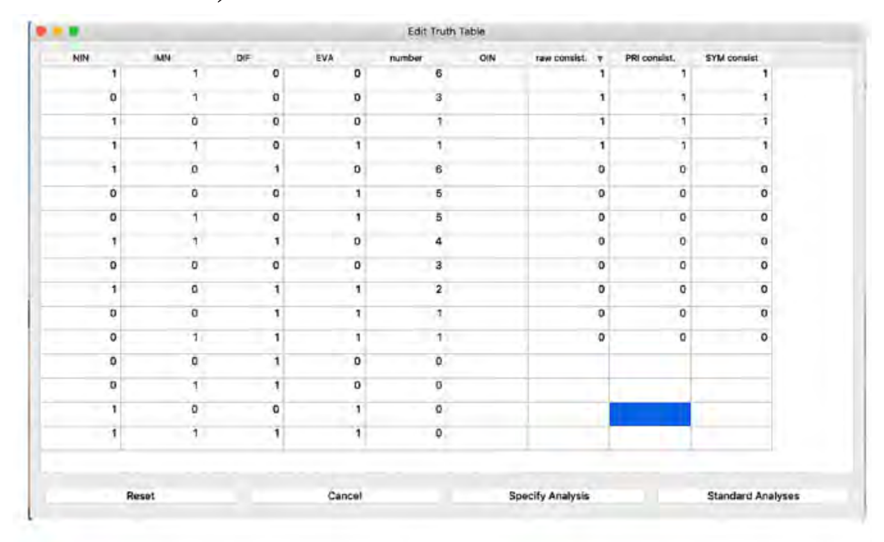



**7.** El siguiente paso es presionar el botón **Visualize**, con lo que se obtiene un *diagrama de Venn* con los datos y las *posibles contradicciones*. En este caso *no hay contradicciones detectadas*. La letra **C**, en la leyenda situada en la parte inferior de la imagen, indica *contradicción*.



## Identificar filas con remanentes lógicos. Caso 4

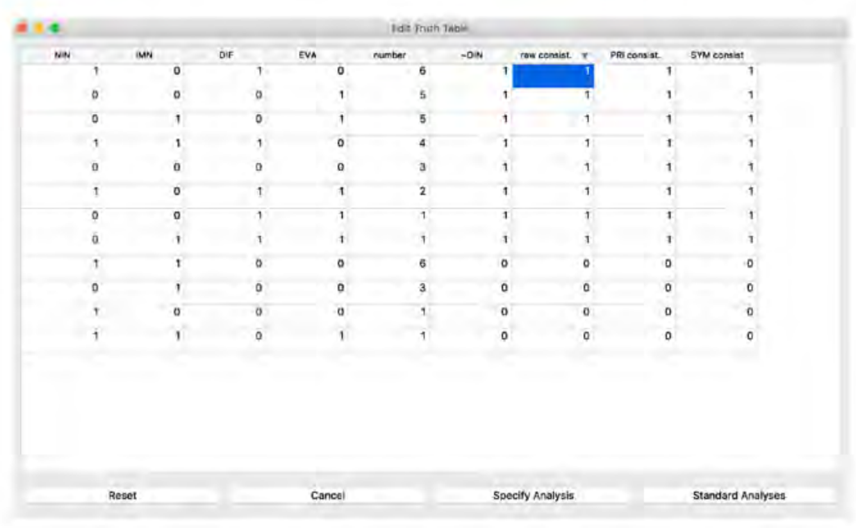
Como ya se mencionó en este capítulo y en anteriores, la manera simple de identificar *residuales lógicos* es a través de la observación de aquellas filas de la *tabla de verdad que no contienen casos*. Si se observan las imágenes de las *tablas de verdad* (**puntos 4 y 11** de la sección *Representar datos en la tabla de verdad*):



Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 3 | | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 | | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 2 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | | | | |
| 0 | 1 | 1 | 0 | 0 | | | | |
| 1 | 0 | 0 | 1 | 0 | | | | |
| 1 | 1 | 1 | 1 | 0 | | | | |

Reset | Cancel | Specify Analysis | Standard Analyses



Edit Truth Table

| NIN | IMN | DIF | EVA | number | ~OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 6 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 1 | 5 | 1 | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 | 5 | 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 4 | 1 | 1 | 1 | 1 |
| 0 | 0 | 0 | 0 | 3 | 1 | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 0 | 6 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |

Reset | Cancel | Specify Analysis | Standard Analyses

Estas presentan cuatro residuales, los mismos tanto para **OIN** como **(~ OIN).** En la *tabla de verdad* se presentan *dicotómicamente* **(1,1,1,1), (1,0,0,1), (0,1,1,0) y (0,0,1,0):**

**(NIN * IMN * DIF * EVA)**

**(NIN * ~ IMN * ~ DIF * EVA)**

**(~NIN * IMN * DIF * ~EVA)**

**(~NIN * ~IMN * DIF * ~EVA)**

El manejo de estos remanentes se describe a continuación en el análisis de suficiencia para los resultados **(1)** y **(0)**.

## Análisis de suficiencia para los resultados (1) y (0). Caso 5

Antes del proceso de *minimización*, se deben analizar las *tablas de verdad* para **OIN** y **(~OIN).** Ya se ha dicho que las filas que presentan configuraciones con una consistencia mayor a **(0.8)** son consideradas como *suficientes* y, por lo tanto, entran en el análisis. En la *tabla de verdad* para **OIN** presentada en la imagen (**punto 4** de la sección *Representar datos en la tabla de verdad*):

Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | | 1 | 1 | 1 |
| 0 | 1 | 0 | 0 | 3 | | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | | 1 | 1 | 1 |
| 1 | 0 | 1 | 0 | 6 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 0 | 1 | 0 | 1 | 5 | | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 4 | | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 3 | | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 | 2 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | | | | |
| 0 | 1 | 1 | 0 | 0 | | | | |
| 1 | 0 | 0 | 1 | 0 | | | | |
| 1 | 1 | 1 | 1 | 0 | | | | |

Reset Cancel Specify Analysis Standard Analyses

Entran en el análisis las filas **(1,2,3,4)**, por tener una consistencia superior a **0.8**, que son las siguientes:

**(NIN * IMN * ~ DIF * ~ EVA)**

**(~NIN * IMN * ~ DIF * ~ EVA)**

**(NIN * IMN * ~ DIF * EVA)**

**(NIN * ~IMN * ~ DIF * ~ EVA)**

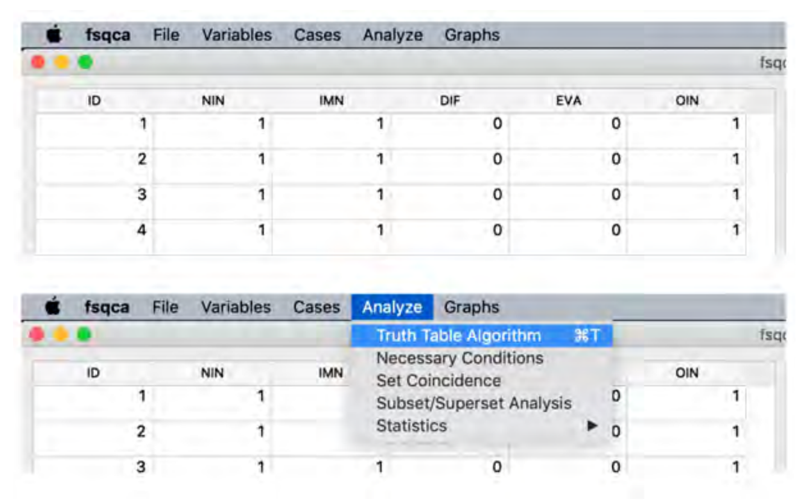
A su vez, los residuales restantes son analizados de acuerdo a la *prueba de necesidad* ejecutado en la *sección análisis de condiciones necesarias*. En este se detectó que **(~ DIF)** es una condición necesaria para **OIN,** por lo tanto los residuales que contengan esta condición deben ser estudiados cuidadosamente, seleccionados e incluidos en el proceso de *minimización* para la obtención de la *solución parsimoniosa e intermedia* y no así para la obtención de la *solución compleja*. En consecuencia, el *residual* de la **fila 14 (NIN * ~IMN * ~ DIF * EVA)** se incluye en el análisis.

Para la *tabla de verdad* referida a **(~ OIN)** el procedimiento es similar. Se puede observar que las ocho primeras configuraciones cumplen con el criterio de *suficiencia* al tener una *consistencia* superior a **0.8**, exactamente de **(1.0)**. Para la obtención de la *solución parsimoniosa e intermedia*, y en relación a los remanentes lógicos, se debe quitar del análisis cualquier remanente que contenga la condición **(~ DIF)**, puesto que es necesaria para la presencia del resultado y no para explicar su *ausencia*. Así, se agrega el valor **(0)** en la columna **(~OIN)** de la configuración **(NIN * ~IMN * ~ DIF * EVA).** Los tres residuales restantes son agregados por los autores para su *minimización*, a saber, **(~NIN * ~IMN * DIF *~EVA), (~NIN * IMN * DIF *~EVA),** y **(NIN * IMN * DIF * EVA),** agregando el valor **(1)** en la fila del resultado.

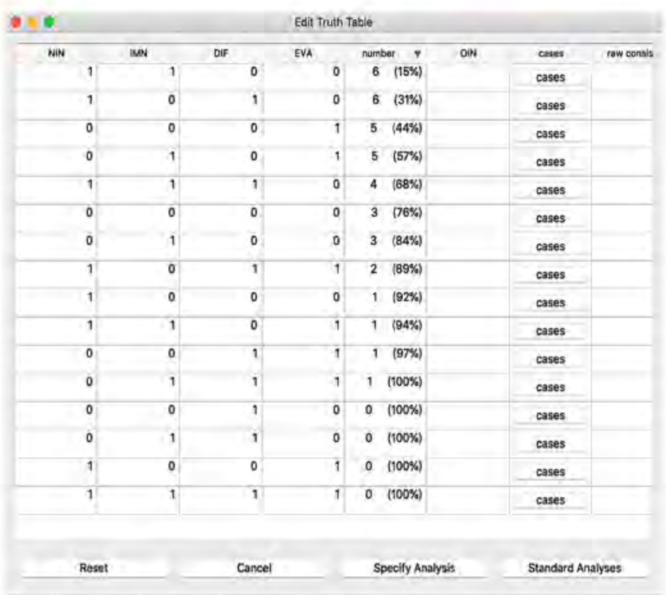
## Minimización de la Tabla de Verdad para la ocurrencia de (OIN). Caso 6

Como se sabe, en su *proceso de minimización*, el software **fsQCA 3.0** provee tres fórmulas diferentes: *compleja*, *parsimoniosa e intermedia*. Cada una de estas fórmulas es obtenida mediante el tipo de tratamiento que hacemos de los residuales lógicos. Para la obtención de la solución compleja de **OIN** se *minimizarán* solamente las configuraciones de los casos observados. Por lo tanto, en el software:

1. Del archivo de trabajo: **values based innovation.csv**, y para realizarlo, en resumen, se analizan los datos dicotómicos a través de una *tabla de verdad* eliminan todas las filas con una consistencia menor a **0.8** y con números inferiores a **(1)**. Al oprimir en la opción OK de la opción **Delete and Code**, la *tabla de verdad* se queda solo con los casos observados como se observa en las imágenes:

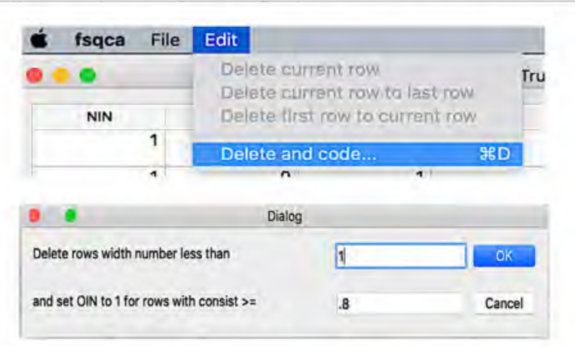


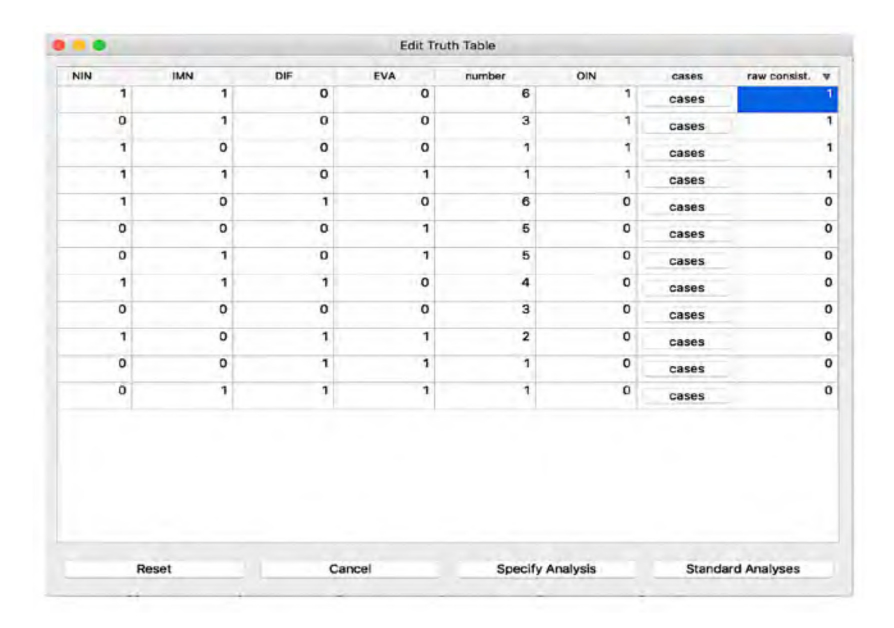




Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | cases | raw consis |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 (15%) | | cases | |
| 1 | 0 | 1 | 0 | 6 (31%) | | cases | |
| 0 | 0 | 0 | 1 | 5 (44%) | | cases | |
| 0 | 1 | 0 | 1 | 5 (57%) | | cases | |
| 1 | 1 | 1 | 0 | 4 (68%) | | cases | |
| 0 | 0 | 0 | 0 | 3 (76%) | | cases | |
| 0 | 1 | 0 | 0 | 3 (84%) | | cases | |
| 1 | 0 | 1 | 1 | 2 (89%) | | cases | |
| 1 | 0 | 0 | 0 | 1 (92%) | | cases | |
| 1 | 1 | 0 | 1 | 1 (94%) | | cases | |
| 0 | 0 | 1 | 1 | 1 (97%) | | cases | |
| 0 | 1 | 1 | 1 | 1 (100%) | | cases | |
| 0 | 0 | 1 | 0 | 0 (100%) | | cases | |
| 0 | 1 | 1 | 0 | 0 (100%) | | cases | |
| 1 | 0 | 0 | 1 | 0 (100%) | | cases | |
| 1 | 1 | 1 | 1 | 0 (100%) | | cases | |

Reset | Cancel | Specify Analysis | Standard Analyses

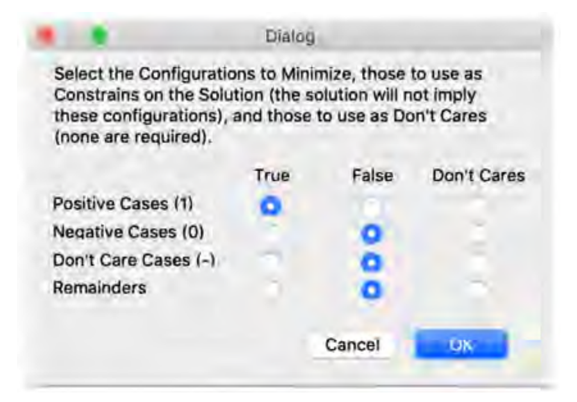



Edit Truth Table

| NIN | IMN | DIF | EVA | number | OIN | cases | raw consist. |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | 1 | cases | 1 |
| 0 | 1 | 0 | 0 | 3 | 1 | cases | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | cases | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 | cases | 1 |
| 1 | 0 | 1 | 0 | 6 | 0 | cases | 0 |
| 0 | 0 | 0 | 1 | 5 | 0 | cases | 0 |
| 0 | 1 | 0 | 1 | 5 | 0 | cases | 0 |
| 1 | 1 | 1 | 0 | 4 | 0 | cases | 0 |
| 0 | 0 | 0 | 0 | 3 | 0 | cases | 0 |
| 1 | 0 | 1 | 1 | 2 | 0 | cases | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | cases | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | cases | 0 |

Reset Cancel Specify Analysis Standard Analyses

2. Después, al presionar en **Specify Analysis**, una nueva ventana es obtenida, la cual nos consulta respecto a cuáles *configuraciones minimizar* para obtener la *solución compleja*, como se muestra:



```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- TRUTH TABLE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
Assumptions:
                    raw         unique
                  coverage     coverage    consistency
                 ----------   ----------   ----------
NIN*~DIF*~EVA     0.636364    0.0909091    1
IMN*~DIF*~EVA     0.818182    0.272727     1
NIN*IMN*~DIF      0.636364    0.0909091    1
solution coverage: 1
solution consistency: 1


Cases with greater than 0.5 membership in term NIN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 11 (1,1)
Cases with greater than 0.5 membership in term IMN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 7 (1,1),
  8 (1,1), 9 (1,1)
Cases with greater than 0.5 membership in term NIN*IMN*~DIF: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 10 (1,1)
```

**3.** Como puede apreciarse, la *solución compleja* es descriptiva y contiene tres de las cuatro condiciones estudiadas para la ocurrencia del fenómeno. La solución en sí está compuesta por *tres posibles rutas* (*equifinalidad*), cada una de las cuales es una combinación de condiciones que deben permanecer juntas para ser *suficientes* y contribuir al *resultado* (conjunción). Otra forma de ver los resultados del reporte es mediante la **Tabla 10.2.**

**Tabla 10.2. Solución *compleja* para la ocurrencia de innovación abierta OIN sin residuales lógicos**

| | **(NIN * ~ DIF * ~ EVA)** | **(IMN * ~ DIF * ~ EVA)** | **(NIN * IMN * ~ DIF)** |
|---|---|---|---|
| **Número de empresas que practican innovación abierta explicadas** | **7** | **9** | **7** |
| **Empresas que practican innovación abierta desglosadas** | **(1, 2, 3, 4, 5, 6, 11)** | **(1, 2, 3, 4, 5, 6, 7, 8, 9)** | **(1, 2, 3, 4, 5, 6, 10)** |
| Raw coverage | **0.636364** | **0.818182** | **0.646364** |
| Unique coverage | **0.0909091** | **0.272727** | **0.636364** |
| Consistency | **1.0** | **1.0** | **1.0** |
| Solution coverage 1.0 | — | — | — |
| Solution consistency 1.0 | — | — | — |

Fuente: **fsQCA 3.0** con datos propios.

**4.** Del mismo modo, se puede apreciar que (~ **DIF)** es necesaria para la ocurrencia del resultado, dado que está incluida en cada parte de la solución. Esto quiere decir que la ocurrencia de **OIN**, no sería posible sin la ausencia de la condición **(~ DIF).**

**5.** Para la obtención de la *solución más parsimoniosa*, se deben incluir los *residuales lógicos* de acuerdo tanto a la evidencia empírica encontrada y a la justificación teórica escogida como a los datos obtenidos del *análisis de necesidad*. En el análisis de necesidad se destacó que **(~ DIF)** es necesaria para **OIN**, por lo tanto cualquier con iguración de *residuales* lógicos que la contenga, debe ser incluida en el *proceso de minimización* y los restantes deben ser removidos del análisis. Los autores han incluido el *residual* **(NIN * ~IMN * ~ DIF * EVA).**

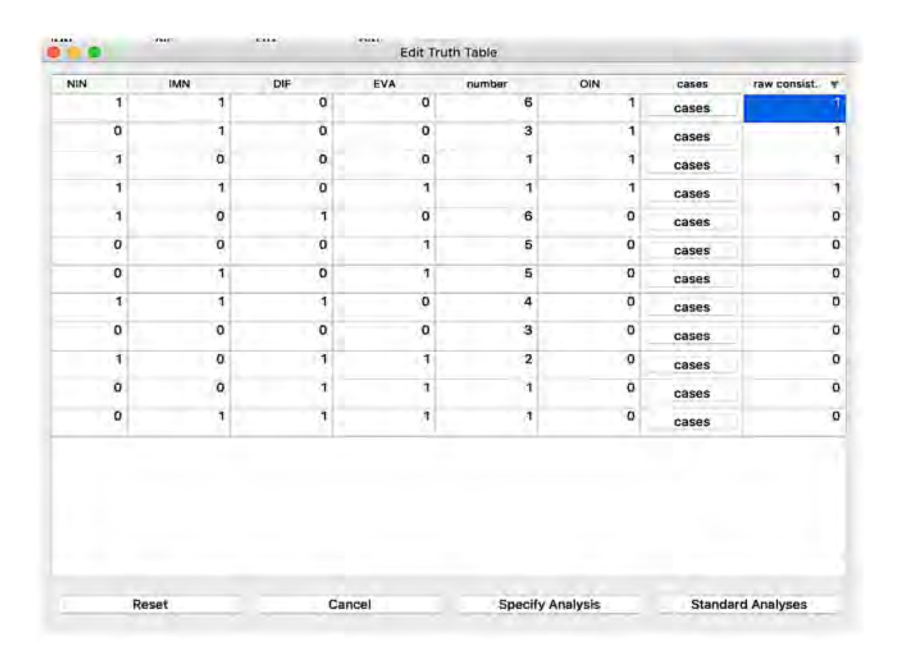
Es aquí, donde interviene otra modalidad llamada **Standard Analyses,** estrategia a desarrollada por **Ragin (2008); Ragin y Sonnet, 2004)** cuando se enfrenta a una *tabla de verdad* que contiene *residuales lógicos.*

Esta estrategia tiene como fin, la reducción de dichos residuales, a través de la producción de *tres soluciones* ejecutadas por el software **fsQCA**:

**a.** *La solución compleja* (donde no existe ninguna supuesto respecto a los *residuales*);
**b.** *La parsimoniosa* (donde se incluyen todas las supuestos simplificadores) y finalmente,
**c.** *La solución intermedia* (donde se incluyen *contrafactuales fáciles*). En este proceso, el investigador es consultado para orientar la *presencia/ausencia* de las condiciones, a través del despliegue de una tabla de expectativas direccionales. Si el investigador cree oportuno que cierta condición debe estar presente para contribuir al resultado, entonces selecciona una determinada condicion en la casilla **present** y si considera que debe estar ausente, pues la selecciona en la respectiva casilla **absent**. Es importante insistir, en que cada decisión debe ser basada en fundamentos teóricos y empíricos **(Schneider y Wagemann, 2012, p. 175)**.

**6.** Por lo tanto, el procedimiento es el siguiente:

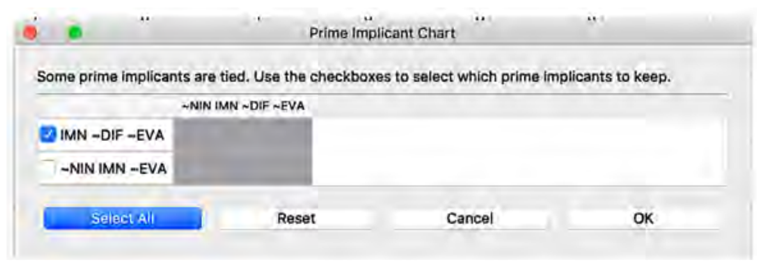
**a.** Posicionarse en la *tabla de verdad* del archivo de trabajo: **value based innovation.csv** en estatus **Delete Code** aplicado y tras haber procedido con el análisis de suficiencia de la *tabla de verdad* descrito anteriormente sobre la ocurrencia del fenómeno **OIN,** se procede a oprimir **Standard Analyses**, que consiste en los siguientes pasos:


Edit Truth Table

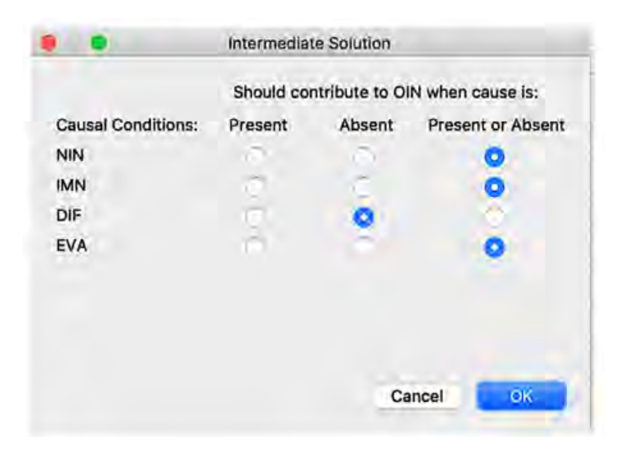
| NIN | IMN | DIF | EVA | number | OIN | cases | raw consist. |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 | 1 | cases | 1 |
| 0 | 1 | 0 | 0 | 3 | 1 | cases | 1 |
| 1 | 0 | 0 | 0 | 1 | 1 | cases | 1 |
| 1 | 1 | 0 | 1 | 1 | 1 | cases | 1 |
| 1 | 0 | 1 | 0 | 6 | 0 | cases | 0 |
| 0 | 0 | 0 | 1 | 5 | 0 | cases | 0 |
| 0 | 1 | 0 | 1 | 5 | 0 | cases | 0 |
| 1 | 1 | 1 | 0 | 4 | 0 | cases | 0 |
| 0 | 0 | 0 | 0 | 3 | 0 | cases | 0 |
| 1 | 0 | 1 | 1 | 2 | 0 | cases | 0 |
| 0 | 0 | 1 | 1 | 1 | 0 | cases | 0 |
| 0 | 1 | 1 | 1 | 1 | 0 | cases | 0 |

Reset Cancel Specify Analysis Standard Analyses

**b.** Para elegir las *implicaciones principales*, el programa emplea un algoritmo que intenta reducir la *tabla de verdad*, hasta que no sea posible una mayor simplificación, comenzando con los implicaciones principales esenciales (las que cubren únicamente, filas específicas en la *tabla de verdad*), que deben aparecer en la solución. Si se ejecuta el algoritmo y la tabla no se puede reducir por completo, el usuario puede seleccionar las implicaciones que se utilizarán, en función de su conocimiento teórico y empírica sustantiva**.** Aparecerá la ventana **Prime Implicant Chart**, que muestra las posibles implicaciones principales para que el usuario las elija. Cada columna en el gráfico, representa una fila diferente de la *tabla de verdad*, que está cubierta por más de una *implicación principal*, como se muestra:



**c.** Una nueva ventana se despliega consultando sobre las expectativas direccionales de las condiciones. De acuerdo con el *análisis de necesidad*, en este ejercicio se escogerá **DIF** como una *condición que debe estar ausente* para la ocurrencia de **OIN**. Para el resto de las opciones se marca la opción **Present or Absent**. A continuación, se presiona en **OK** para obtener la *solución más parsimoniosa* y la *solución intermedia.*




**d.** Se obtienen, tres resultados a verificar:

**Solución compleja** (observe que arroja los mismos resultados que cuando se hizo con la opción: **Specify Analysis**):

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
                    raw         unique
                    coverage    coverage    consistency
                    ----------  ----------  ----------
NIN*~DIF*~EVA       0.636364    0.0909091   1
IMN*~DIF*~EVA       0.818182    0.272727    1
NIN*IMN*~DIF        0.636364    0.0909091   1
solution coverage: 1
solution consistency: 1
```

```
Cases with greater than 0.5 membership in term NIN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 11 (1,1)
Cases with greater than 0.5 membership in term IMN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 7 (1,1),
  8 (1,1), 9 (1,1)
Cases with greater than 0.5 membership in term NIN*IMN*~DIF: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 10 (1,1)
```

La solución **parsimoniosa:**

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
                      raw        unique
                    coverage    coverage    consistency
                   ----------  ----------  ----------
NIN*~DIF           0.727273    0.181818    1
IMN*~DIF*~EVA      0.818182    0.272727    1
solution coverage: 1
solution consistency: 1


Cases with greater than 0.5 membership in term NIN*~DIF: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 10 (1,1),
  11 (1,1)
Cases with greater than 0.5 membership in term IMN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 7 (1,1),
  8 (1,1), 9 (1,1)
```

Contiene *dos rutas* (*equifinalidad*) que en sí son suficientes para explicar la ocurrencia de **OIN.** Cada ruta es una *conjunción* de condiciones que contiene la ausencia de **(~ DIF)** como necesaria para el resultado. Si **(~ DIF)** no hubiera sido señalada como necesaria en el análisis de necesidad, entonces este sería una condición **INUS** en las soluciones.

Así mismo, se presenta la **solución intermedia***:*

```
***********************
*TRUTH TABLE ANALYSIS*
***********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
Assumptions:
~DIF (absent)
                    raw        unique
                  coverage    coverage    consistency
                 ----------  ----------  ----------
NIN*~DIF*~EVA     0.636364    0.0909091   1
IMN*~DIF*~EVA     0.818182    0.272727    1
NIN*IMN*~DIF      0.636364    0.0909091   1
solution coverage: 1
solution consistency: 1


Cases with greater than 0.5 membership in term NIN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 11 (1,1)
Cases with greater than 0.5 membership in term IMN*~DIF*~EVA: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 7 (1,1),
  8 (1,1), 9 (1,1)
Cases with greater than 0.5 membership in term NIN*IMN*~DIF: 1 (1,1),
  2 (1,1), 3 (1,1), 4 (1,1),
  5 (1,1), 6 (1,1), 10 (1,1)
```

**e.** Otra forma de ver los resultados del reporte es mediante las **tablas 10.3 y 10.4.**

**Tabla 10.3. Solución *más parsimoniosa* para la ocurrencia de innovación abierta, OIN**

| | **(NIN * ~ DIF)** | **(IMN * ~ DIF * ~ EVA)** |
|---|---|---|
| **Número de empresas que practican innovación abiertaexplicadas** | **8** | **9** |
| **Empresas que practican innovación abiertadesglosadas** | **(1, 2, 3, 4, 5, 6, 10, 11)** | **(1, 2, 3, 4, 5, 6, 7, 8, 9)** |
| Raw coverage | **0.727273** | **0.818182** |
| Unique coverage | **0.181818** | **0.272727** |
| Consistency | **1.0** | **1.0** |
| Solution coverage 1.0 | — | — |
| Solution consistency 1.0 | — | — |

Fuente: **fsQCA 3.0** con datos propios.

**Tabla 10.4. Solución *intermedia* para la ocurrencia de innovación abierta, OIN**

| | **(NIN * ~ DIF* ~EVA)** | **(IMN * ~ DIF * ~EVA)** | **(NIN * IMN * ~DIF*)** |
|---|---|---|---|
| **Número de empresas que practican innovación abierta explicadas** | **7** | **9** | **7** |
| **Empresas que practican innovación abierta desglosadas** | **(1, 2, 3, 4, 5, 6, 10, 11)** | **(1, 2, 3, 4, 5, 6, 7, 8, 9)** | **(1, 2, 3, 4, 5, 6,10)** |
| Raw coverage | **0.636364** | **0.818182** | **0.636364** |
| Unique coverage | **0.0909091** | **0.272727** | **0.0909091** |
| Consistency | **1.0** | **1.0** | **1.0** |
| Solution coverage 1.0 | — | — | — |
| Solution consistency 1.0 | — | — | — |

Fuente: **fsQCA 3.0** con datos propios.

**f.** Al comparar la solución más *parsimoniosa* con la *intermedia*, esta última es la recomendable a utilizar por considerar más parámetros para su explicación, de tal forma que la solución, se conforma por la combinación de tres causalidades para obtener el resultado de interés, es:

**(NIN * ~ DIF* ~EVA) + (IMN * ~ DIF* ~EVA)**
**+ (NIN * IMN * ~DIF*) → OIN**

**g.** La narrativa de la solución deberá leerse de forma tal que las conjunciones y las uniones por grupos le den claridad a la explicación que se describa, por ejemplo:

"*Las empresas que practican innovación abierta tienen una alta efectividad* **OIN** *cuando existe alta innovación normativa* **NIN** ***aunque*** *no tengan diferenciación en la innovación del producto/servicio* **(~ DIF)** *y* ***aunque*** *no tengan un modelo de evaluación de valores* **(~ EVA)**, ***tal y como ocurre*** *con las empresas* **(1,2,3,4,5,6,11) o cuando** *tienen una alta adopción de los modelos de negocios* **IMN** ***aunque*** *no tengan diferenciación en la innovación del producto/servicio* **(~ DIF)** *y* ***aunque*** *no tengan un modelo de evaluación de valores* **(~ EVA)**, ***tal y como ocurre*** *con las empresas* **(1, 2, 3, 4, 5, 6, 7, 8, 9), o cuando** *existen con alta innovación normativa* **(NIN)** *con una alta adopción de los modelos de negocios* **(IMN)** ***aunque*** *no tengan diferenciación en la innovación del producto/servicio* **(~ DIF)**, *tal y como ocurre con las empresas* **(1, 2, 3, 4, 5, 6, 10)**".

**h.** En una reducción de la expresión, a:

**[(NIN + IMN) * (~ DIF* ~EVA)] + (NIN * IMN * ~DIF*) → OIN**

Se pudiera redactar más concretamente a:

"*Las empresas que practican innovación abierta tienen una alta efectividad* **OIN** *cuando existe alta innovación normativa* **NIN** ***o*** *una alta adopción de los modelos de negocios* **IMN** ***a pesar*** *de que no tengan diferenciación en la innovación del producto/servicio*

**(~ DIF) y** *no tengan no tengan un modelo de evaluación de valores* **(~ EVA),** ***tal y como ocurre*** *con las empresas* **(1, 2, 3, 4, 5, 6, 7, 8, 9, 11)** ***o cuando*** *existe alta innovación normativa* **NIN** ***con*** *una alta adopción de los modelos de negocios* **IMN** ***aunque no*** *no tengan diferenciación en la innovación del producto/servicio* **(~ DIF),** ***tal y como ocurre*** *con las empresas* **(1, 2, 3, 4, 5, 6, 10)**".

Es recomendable elegir la opción que más apegado se encuentre tanto al marco teórico de la investigación así como al conocimiento empírico y sustantivo, del investigador, debiendo considerar complementariamente, los valores de *consistencia* y *cobertura* de cada vía, tanto de la solución *parsimoniosa, como la intermedia* por ejemplo, se puede observar que la cobertura de la segunda es mayor que en la primera. Del mismo modo, la consistencia de la solución con sus rutas que son de **(1.0)** y su cobertura cubre el **100%** de los casos.

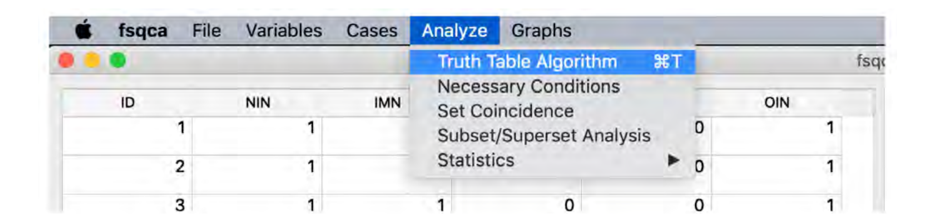
## Minimización de la Tabla de Verdad para la ausencia de (~ OIN). Caso 7

Para las configuraciones de **(~ OIN)** se debe:

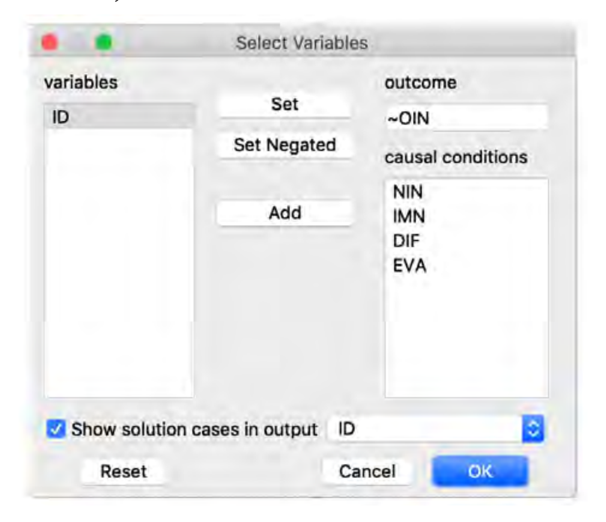
1. Posicionar en la *tabla de verdad* del archivo de trabajo: **value based innovation.csv** en estatus **Delete Code** aplicado y tras haber procedido con el análisis de suficiencia de la *tabla de verdad* descrito anteriormente sobre la ocurrencia del fenómeno **OIN,** se procede a oprimir **Standard Analyses**, que consiste en los siguientes pasos:

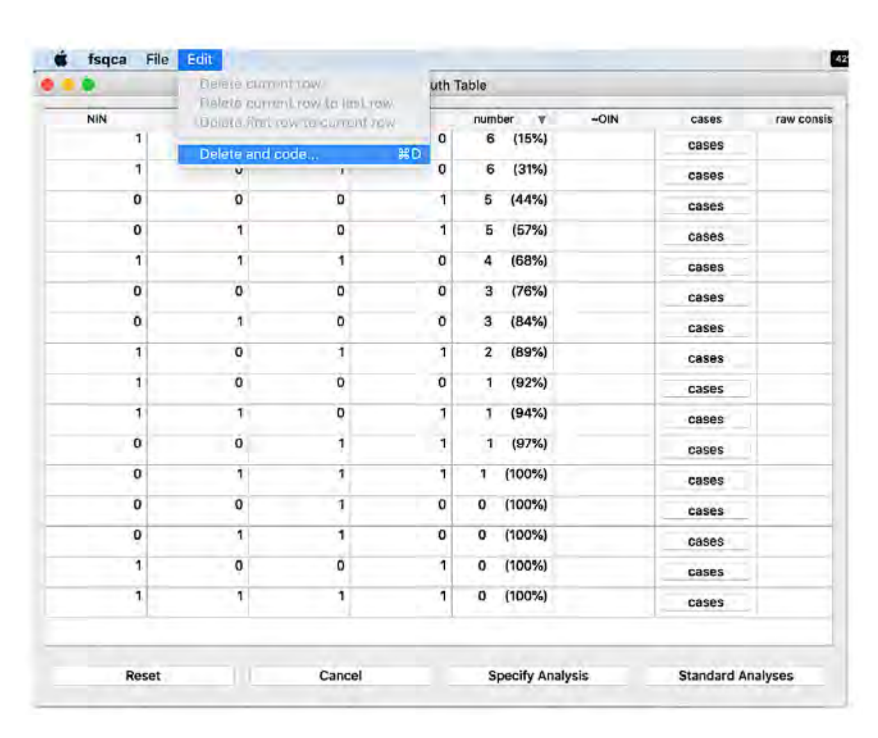
fsqca File Variables Cases Analyze Graphs

| ID | NIN | IMN | DIF | EVA | OIN |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 0 | 1 |
| 2 | 1 | 1 | 0 | 0 | 1 |
| 3 | 1 | 1 | 0 | 0 | 1 |
| 4 | 1 | 1 | 0 | 0 | 1 |

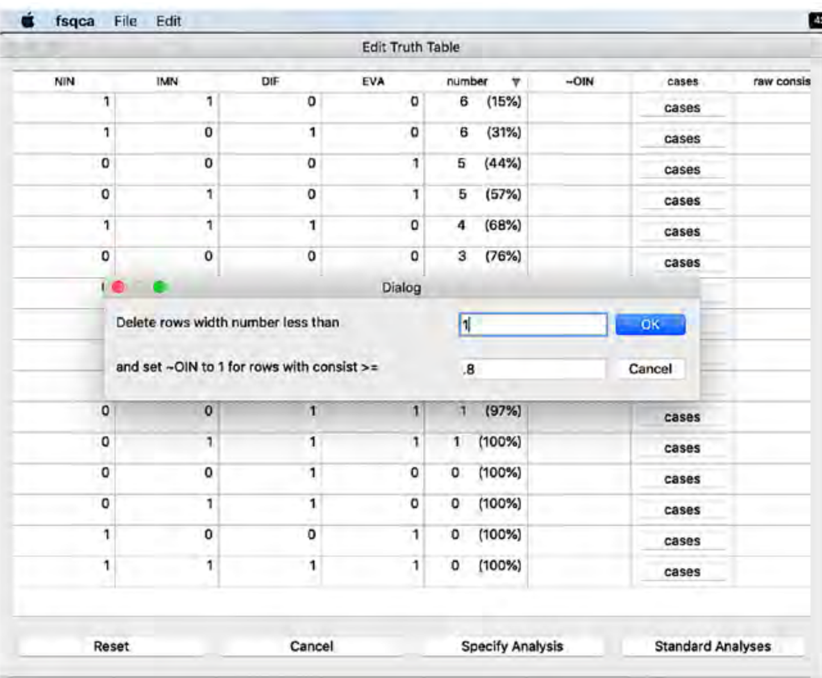



**2.** Incluir el resultado en la casilla de **Outcome** oprimiendo **Set Negated** agregando todas las condiciones y proceder, como se ha explicado en apartados anteriores, con la definición de los criterios de *suficiencia* en la *tabla de verdad*, como se muestra:

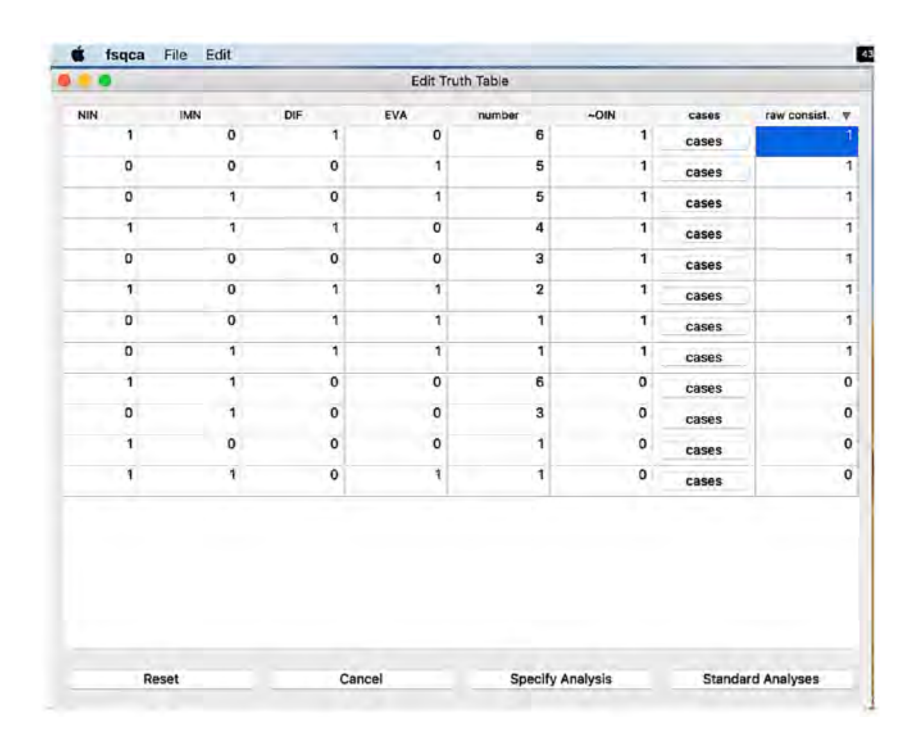





| NIN | | | | number | ~OIN | cases | raw consis |
|---|---|---|---|---|---|---|---|
| 1 | | | 0 | 6 (15%) | | cases | |
| 1 | | | 0 | 6 (31%) | | cases | |
| 0 | 0 | 0 | 1 | 5 (44%) | | cases | |
| 0 | 1 | 0 | 1 | 5 (57%) | | cases | |
| 1 | 1 | 1 | 0 | 4 (68%) | | cases | |
| 0 | 0 | 0 | 0 | 3 (76%) | | cases | |
| 0 | 1 | 0 | 0 | 3 (84%) | | cases | |
| 1 | 0 | 1 | 1 | 2 (89%) | | cases | |
| 1 | 0 | 0 | 0 | 1 (92%) | | cases | |
| 1 | 1 | 0 | 1 | 1 (94%) | | cases | |
| 0 | 0 | 1 | 1 | 1 (97%) | | cases | |
| 0 | 1 | 1 | 1 | 1 (100%) | | cases | |
| 0 | 0 | 1 | 0 | 0 (100%) | | cases | |
| 0 | 1 | 1 | 0 | 0 (100%) | | cases | |
| 1 | 0 | 0 | 1 | 0 (100%) | | cases | |
| 1 | 1 | 1 | 1 | 0 (100%) | | cases | |

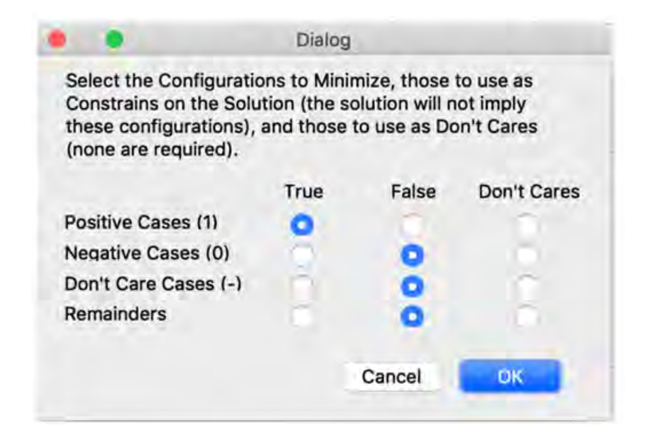


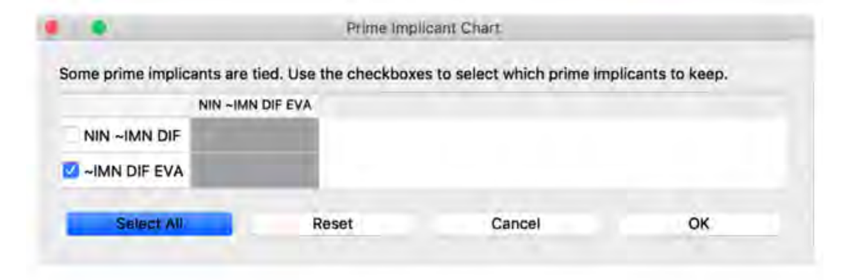

| NIN | IMN | DIF | EVA | number | ~OIN | cases | raw consis |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 6 (15%) | | cases | |
| 1 | 0 | 1 | 0 | 6 (31%) | | cases | |
| 0 | 0 | 0 | 1 | 5 (44%) | | cases | |
| 0 | 1 | 0 | 1 | 5 (57%) | | cases | |
| 1 | 1 | 1 | 0 | 4 (68%) | | cases | |
| 0 | 0 | 0 | 0 | 3 (76%) | | cases | |
| 0 | 0 | 1 | 1 | 1 (97%) | | cases | |
| 0 | 1 | 1 | 1 | 1 (100%) | | cases | |
| 0 | 0 | 1 | 0 | 0 (100%) | | cases | |
| 0 | 1 | 1 | 0 | 0 (100%) | | cases | |
| 1 | 0 | 0 | 1 | 0 (100%) | | cases | |
| 1 | 1 | 1 | 1 | 0 (100%) | | cases | |




| NIN | IMN | DIF | EVA | number | ~OIN | cases | raw consist. |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 6 | 1 | cases | 1 |
| 0 | 0 | 0 | 1 | 5 | 1 | cases | 1 |
| 0 | 1 | 0 | 1 | 5 | 1 | cases | 1 |
| 1 | 1 | 1 | 0 | 4 | 1 | cases | 1 |
| 0 | 0 | 0 | 0 | 3 | 1 | cases | 1 |
| 1 | 0 | 1 | 1 | 2 | 1 | cases | 1 |
| 0 | 0 | 1 | 1 | 1 | 1 | cases | 1 |
| 0 | 1 | 1 | 1 | 1 | 1 | cases | 1 |
| 1 | 1 | 0 | 0 | 6 | 0 | cases | 0 |
| 0 | 1 | 0 | 0 | 3 | 0 | cases | 0 |
| 1 | 0 | 0 | 0 | 1 | 0 | cases | 0 |
| 1 | 1 | 0 | 1 | 1 | 0 | cases | 0 |

**3.** A partir de ello, para obtener la **solución compleja** se deben aplicar los *criterios de su iciencia*, y dejar tan solo los casos observados *sin incluir los residuales lógicos*. Se procede entonces a *minimizar* las con igu-raciones mediante la opción **Specify Analysis**. Una ventana se abrirá solicitando seleccionar la *con iguración a minimizar*: Además de una implicación principal a seleccionar, en nuestro caso será **(-IMN * DIF * EVA)**. Seleccionar como se muestra y presionar **OK**:




```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: ~OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- TRUTH TABLE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
Assumptions:
                    raw         unique
                  coverage     coverage    consistency
                 ----------   ----------   ----------
~NIN*EVA         0.444444     0.222222     1
~NIN*~IMN*~DIF   0.296296     0.111111     1
NIN*DIF*~EVA     0.37037      0.37037      1
~IMN*DIF*EVA     0.111111     0.0740741    1
solution coverage: 1
solution consistency: 1
```

```
Cases with greater than 0.5 membership in term ~NIN*EVA: 25 (1,1),
  26 (1,1), 27 (1,1), 28 (1,1),
  29 (1,1), 30 (1,1), 31 (1,1),
  32 (1,1), 33 (1,1), 34 (1,1),
  37 (1,1), 38 (1,1)
Cases with greater than 0.5 membership in term ~NIN*~IMN*~DIF: 22 (1,1),
  23 (1,1), 24 (1,1), 30 (1,1),
  31 (1,1), 32 (1,1), 33 (1,1),
  34 (1,1)
Cases with greater than 0.5 membership in term NIN*DIF*~EVA: 12 (1,1),
  13 (1,1), 14 (1,1), 15 (1,1),
  16 (1,1), 17 (1,1), 18 (1,1),
  19 (1,1), 20 (1,1), 21 (1,1)
Cases with greater than 0.5 membership in term ~IMN*DIF*EVA: 35 (1,1),
  36 (1,1), 38 (1,1)
```

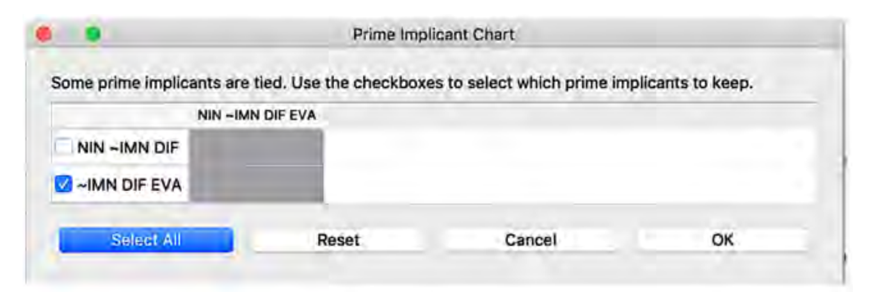
**4.** Como puede apreciarse, la **solución compleja** es descriptiva y contiene cuatro condiciones estudiadas para la ocurrencia del fenómeno. La solu-ción en sí está compuesta por *cuatro posibles rutas* (*equifinalidad*), cada una de las cuales es una combinación de condiciones que deben perma-necer juntas para ser *su icientes* y contribuir al *resultado* (conjunción). Otra forma de ver los resultados del reporte es mediante la **Tabla 10.5**, donde la opción **(~ NIN * EVA)** cubre un **44%** de los casos.

**Tabla 10.5. Solución *compleja* para la no ocurrencia de innovación abierta OIN sin residuales lógicos**

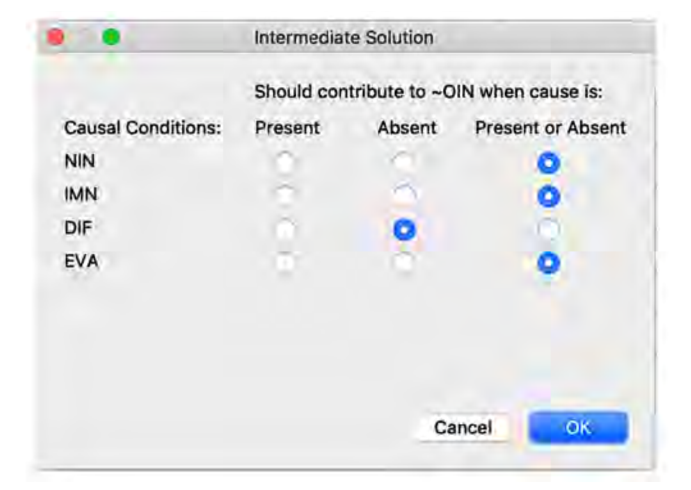
| | (~ NIN * EVA) | (~ NIN * ~ IMN * ~DIF) | (NIN * DIF * ~EVA) | (~ IMN * DIF * EVA) |
|---|---|---|---|---|
| **Número de empresas que practican innovación abierta explicadas** | **12** | **8** | **10** | **3** |
| **Empresas que practican innovación abierta desglosadas** | **(25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 37, 38)** | **(22, 23,24, 30, 31, 32, 33, 34)** | **(12, 13, 14, 15, 16, 17, 18, 19, 20, 21)** | **(35, 36, 38)** |
| Raw coverage | **0.444444** | **0.296296** | **0.37037** | **0.111111** |
| Unique coverage | **0.222222** | **0.111111** | **0.37037** | **0.0740741** |
| Consistency | **1.0** | **1.0** | **1.0** | **1.0** |
| Solution coverage 1.0 | - | - | - | - |
| Solution consistency 1.0 | - | - | - | - |

Fuente: **fsQCA 3.0** con datos propios.

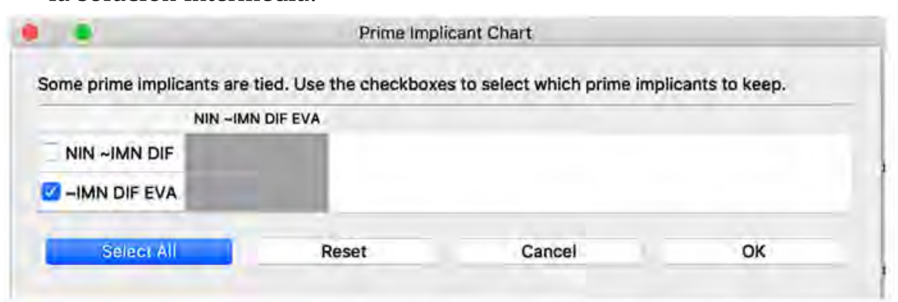
5. Para obtener las opciones **parsimoniosa e intermedia**, habrá que regresar a las condiciones del archivo descritas en el **punto 2** y elegir la opción: **Standard Analyses**. El software **fsQCA 3.0** volverá a preguntar sobre la selección de las implicaciones. Es importante destacar que la opción de las *implicaciones* depende de la teoría y la evidencia empírica que maneje el investigador. La solución compleja de la *ausencia efectividad en la innovación abierta* **(~ OIN)** está compuesta por tres rutas, cada una de las cuales es una conjunción de condiciones que conjuntamente son necesarias para producir el resultado y por lo que se deberá elegir la implicación principal **(~IMN * DIF * EVA)** para la *minimización* y oprimir **OK**:




6. Para la obtención de la **solución parsimoniosa**, se deben incluir aquellos residuales lógicos que cumplan con los *criterios de necesidad* y *suficiencia* y excluir los que no los cumplan. Si bien es cierto, el *análisis de necesidad* no arrojó resultados sobre la existencia de condiciones necesarias para **(~ OIN)** sí se tiene la información sobre la condición necesaria para la ocurrencia de **OIN**. Por lo tanto, cualquier residual que contenga **(~ DIF)** debe ser *excluido del análisis de* **(~ OIN)** con el fin de *no violar el criterio de necesidad y evitar contradicciones*. Para obtener la solución más parsimoniosa el autor incluyó los tres residuales restantes en la minimización. Como se aprecia:




**7.** Vuelve a solicitar confirmar las implicaciones principales para calcular la solución intermedia:




**8.** Observe nuevamente que la **solución compleja** a través de la opción **Standard Analyses** es igual a la que emite la opción **Speci ic Analysis** detallada en el **punto 3** anterior.

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: ~OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
```

```
                     raw          unique
                   coverage      coverage     consistency
                  ----------    ----------    ----------
~NIN*EVA           0.444444     0.222222      1
~NIN*~IMN*~DIF     0.296296     0.111111      1
NIN*DIF*~EVA       0.37037      0.37037       1
~IMN*DIF*EVA       0.111111     0.0740741     1
solution coverage: 1
solution consistency: 1


Cases with greater than 0.5 membership in term ~NIN*EVA: 25 (1,1),
  26 (1,1), 27 (1,1), 28 (1,1),
  29 (1,1), 30 (1,1), 31 (1,1),
  32 (1,1), 33 (1,1), 34 (1,1),
  37 (1,1), 38 (1,1)
Cases with greater than 0.5 membership in term ~NIN*~IMN*~DIF: 22 (1,1),
  23 (1,1), 24 (1,1), 30 (1,1),
  31 (1,1), 32 (1,1), 33 (1,1),
  34 (1,1)
Cases with greater than 0.5 membership in term NIN*DIF*~EVA: 12 (1,1),
  13 (1,1), 14 (1,1), 15 (1,1),
  16 (1,1), 17 (1,1), 18 (1,1),
  19 (1,1), 20 (1,1), 21 (1,1)
Cases with greater than 0.5 membership in term ~IMN*DIF*EVA: 35 (1,1),
  36 (1,1), 38 (1,1)
```

9. La **solución parsimoniosa** presenta tres rutas para la no ocurrencia de innovación abierta **(~ OIN)** .Dos de ellos son conjunciones de condiciones que contribuyen al resultado y tan solo una es una condición inrdividual suficiente para el mismo **DIF**. Así, su presencia sería suficiente para la **(~OIN)**.

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: ~OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
                   raw         unique
                 coverage     coverage     consistency
                ----------   ----------    ----------
DIF              0.518519     0.444444     1
~NIN*~IMN        0.333333     0.111111     1
~NIN*EVA         0.444444     0.185185     1
solution coverage: 1
solution consistency: 1
```

```
Cases with greater than 0.5 membership in term DIF: 12 (1,1),
   13 (1,1), 14 (1,1), 15 (1,1),
   16 (1,1), 17 (1,1), 18 (1,1),
   19 (1,1), 20 (1,1), 21 (1,1),
   35 (1,1), 36 (1,1), 37 (1,1),
   38 (1,1)
Cases with greater than 0.5 membership in term ~NIN*~IMN: 22 (1,1),
   23 (1,1), 24 (1,1), 30 (1,1),
   31 (1,1), 32 (1,1), 33 (1,1),
   34 (1,1), 38 (1,1)
Cases with greater than 0.5 membership in term ~NIN*EVA: 25 (1,1),
   26 (1,1), 27 (1,1), 28 (1,1),
   29 (1,1), 30 (1,1), 31 (1,1),
   32 (1,1), 33 (1,1), 34 (1,1),
   37 (1,1), 38 (1,1)
```

**10.** Finalmente, para obtener la **solución intermedia** debemos realizar ciertos supuestos con respecto a las condiciones incluidas en el estudio. Si aplicamos la Ley de De Morgan, se debería seleccionar la presencia de **DIF** como una condición suficiente para **(~ OIN)**, cuando se oprimió la configuración en la opción **Absent** para la condición **(~ DIF)** y seleccionando **Present or Absent** para las restantes, como se hizo en el punto 6 anterior.

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/value based innovation.csv
Model: ~OIN = f(NIN, IMN, DIF, EVA)
Algorithm: Quine-McCluskey

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 1
Assumptions:
~DIF (absent)
                      raw        unique
                    coverage    coverage   consistency
                   ----------  ----------  ----------
~NIN*EVA            0.444444    0.222222    1
~NIN*~IMN*~DIF      0.296296    0.111111    1
NIN*DIF*~EVA        0.37037     0.37037     1
~IMN*DIF*EVA        0.111111    0.0740741   1
solution coverage: 1
solution consistency: 1


Cases with greater than 0.5 membership in term ~NIN*EVA: 25 (1,1),
  26 (1,1), 27 (1,1), 28 (1,1),
  29 (1,1), 30 (1,1), 31 (1,1),
  32 (1,1), 33 (1,1), 34 (1,1),
  37 (1,1), 38 (1,1)
```

```
Cases with greater than 0.5 membership in term -NIN*-IMN*-DIF: 22 (1,1),
   23 (1,1), 24 (1,1), 30 (1,1),
   31 (1,1), 32 (1,1), 33 (1,1),
   34 (1,1)
Cases with greater than 0.5 membership in term NIN*DIF*-EVA: 12 (1,1),
   13 (1,1), 14 (1,1), 15 (1,1),
   16 (1,1), 17 (1,1), 18 (1,1),
   19 (1,1), 20 (1,1), 21 (1,1)
Cases with greater than 0.5 membership in term -IMN*DIF*EVA: 35 (1,1),
   36 (1,1), 38 (1,1)
```

**11.** Como se puede apreciar, la **solución intermedia** es idéntica a la **compleja**. Es una buena práctica en el **QCA** analizar la **solución intermedia** debido a que se basa en expectativas acerca de las condiciones y es siempre un *subconjunto de la* **solución parsimoniosa** y un *super-conjunto de la* **solución compleja (Schneider y Wagemann, 2012, p.206**

**12.** Las **soluciones intermedia y compleja**, contienen cuatro vías que ex- plican el resultado de interés. La **solución parsimoniosa** tiene tres vías, las cuales son menos complicadas y más claras para llegar al resultado  inal:

**(~ NIN * EVA) + (~NIN * ~IMN * ~DIF) + (NIN * DIF * ~EVA) + (~IMN * DIF * EVA) → (~OIN)**

La narrativa de la solución puede leerse del siguiente modo:

"*Las empresas que practican innovación abierta*, dejan de tenerla **(~OIN)** cuando dejan de ejercer *innovación normativa* **(~ NIN) pero** tienen un modelo de *evaluación de valores* **EVA, tal como ocurre** en las empresas **(25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 37, 38) o** cuando dejan de ejercer *innovación normativa* **(~ NIN) y** dejan de tener *innovación en el modelo de negocios* **(~ IMN) incluyendo** no tener *innovación en el producto/servicio diferenciado* **(~ DIF) tal como ocurre** en las empresas **(22, 23, 24, 30, 31, 32, 33, 34) o** cuando ejercen *innovación normativa* **NIN y** cuentan con *innovación en el producto/servicio diferenciado* **DIF, aún y** cuando no tengan un *modelo de evaluación de valores* **(~ EVA)**, **tal como ocurre** en las empresas **(12, 13, 14, 15, 16, 17, 18,19, 20, 21) o** dejan de tener *innovación en el modelo de negocios*

**(~ IMN) y** cuentan con *innovación en el producto/servicio diferenciado* **DIF y** tengan un *modelo de evaluación de valores* **EVA**, **tal y como ocurre** con las empresas **(35, 36, 38)”.**

**13.** Al observar los datos de *consistencia y cobertura* de cada vía de la solución parsimoniosa es mayor que las compleja e intermedia, se puede observar que la cobertura de uno de los factores más altos de la primera la primera **(DIF)** tiene un **44%,** la cual es mayor que las dos restantes **(NIN * DIF * ~ EVA)** con un **37%.**La solución en sí misma tiene una consistencia y cobertura máxima.

## Interpretación de los resultados. Caso 8

Los resultados se pueden interpretar sobre la base de las expresiones minimizadas o partes de una solución que cumpla con el umbral *preestablecido de consistencia.* Esto es debido a que la consistencia total de la solución es me nos relevante que la de sus partes **(Schneider y Wagemann, 2012, p. 281)**. Para ello, existen dos procesos, como explica **Ragin (1987)**:

- Uno es el retorno a los casos y el otro es la evaluación de la teoría. Ambos procesos, cuando se unen, deben establecer mecanismos causales o explicativos.
- En el ejemplo de ocurrencia en la efectividad de la innovación abierta **OIN**, uno podría preguntarse sobre la narrativa que hay detrás del hecho que, de acuerdo con una expresión tan compleja, discutirla con el conocimiento empírico y buscar su validación a través del marco teórico realizado.

CAPÍTULO 11.

# Guía de análisis fsQCA

Siendo el software **fsQCA 3.0** el más popular para análisis de este tipo, los usuarios familiarizados con **STATA** y **R** bien pueden recurrir a los respectivos paquetes para obtener, con plenas garantías, un análisis *difuso*. Por popularidad y funcionalidad, aquí, se explican los pasos a seguir para hacer un análisis de conjuntos difusos con **fsQCA**. En términos generales, a los habituados al análisis **csQCA**, proceder con un análisis **fsQCA** no les resultará ni mucho menos complicado, salvo en lo referido a consultar **XY Plots**, *calibrar valores* o tener en cuenta los *parámetros de consistencia y cobertura* durante los análisis de condiciones *necesarias y suficientes*. Por lo tanto, es muy importante realizar la *matriz de datos* con capture de la lógica *difusa*.

A partir de entonces, el procedimiento es muy similar al que se tiene con **csQCA**:

**a.** Se solicita la selección de las condiciones causales y del resultado (*en tanto que ocurrencia o no ocurrencia*);
**b.** Se despliega la *tabla de verdad* en la que se deben indicar los casos *positivos* y los casos *negativos* dependiendo de los parámetros de consistencia, **PRI** de cada expresión;
**c.** Se da la posibilidad de destacar expectativas direccionales para producir *soluciones intermedias en el análisis de condiciones suficientes*, que acompañarán a las **soluciones complejas y más parsimoniosas;**
**d.** Se encuentra la opción de analizar las condiciones necesarias, etcétera.

A fin de lograr una explicación más detallada, se expondrán los casos hipotéticos contenidos en el archivo de trabajo: **innovacionmn.cvs,** que muestra la relación de factores que existen en un listado de **14** empresas de innovación en productos de alta tecnología que han adaptado los factores

que se involucran en la *innovación por el modelo de negocios*, acorde a la **OECD (2018),** el cual involucra las etapas de: producción, logística, mercadotecnia y arreglos cooperativos, entre otras, siendo:

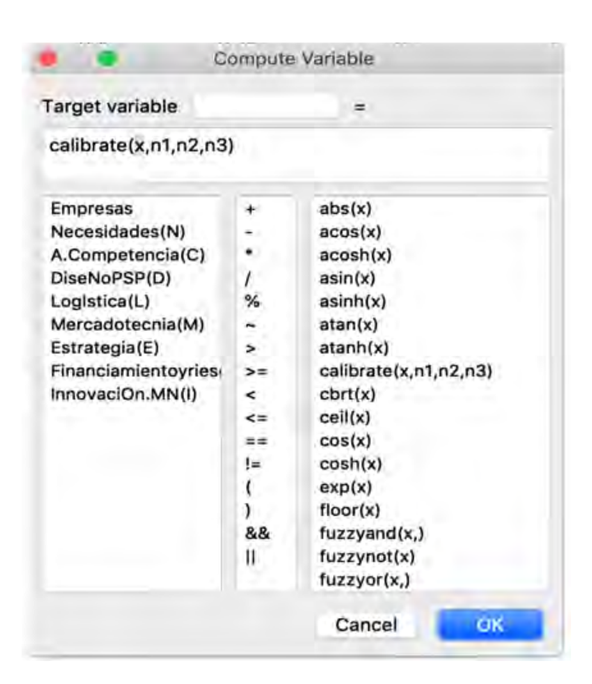
- *Grado de necesidades detectadas del consumidor* **N,** que son susceptibles a proponer nuevos diseños de productos innovadores, especialmente, en lo referente a su simpatía por adoptarlos de manera inmediata.
- *Análisis de la competencia* **C,** como nivel del conocimiento que se tiene de la misma, cuidando que la propuesta de productos innovadores estableciera diferenciales claramente notorios y no esforzarse en productos similares al la competencia, sino más bien, superiores. Precisión de rivales que no lo son en primera instancia.
- *Nivel de diseño del producto de alta tecnología a innovar* **D,** con valor agregado, pruebas prototipo y producción. Etapa que considera los requerimientos tanto del cliente así como las capacidades de desarrollo y fabricación de la empresa para la inversión en prototipos, que permitan hacer la escala a producción masiva.
- *Nivel de recursos de logística* **L,** comprometidos con base a la estrategia de penetración de mercado tanto local, internacional o mundial. Detalla el proceso de requerimiento, distribución y transporte tanto de los insumos así como del producto innovador fabricado.
- *Mercadotecnia* **M**, que incluye la implementación de las innovaciones por mercadotecnia digital, para elevar la presencia en el mercado del producto innovador y mejorar los canales de logística del producto innovador.
- *Estrategia,* **A.** La Estrategia de negocios que corresponde a asociaciones, alianzas, etcétera; que apoyan a una, varias o/a todas las etapas de la innovación del modelo de negocios. Incluye la incorporación de más modelos de negocios relacionados.
- *Financiamiento y riesgo* **F.** Análisis financiero de los recursos implicados con el riesgo que se corre.

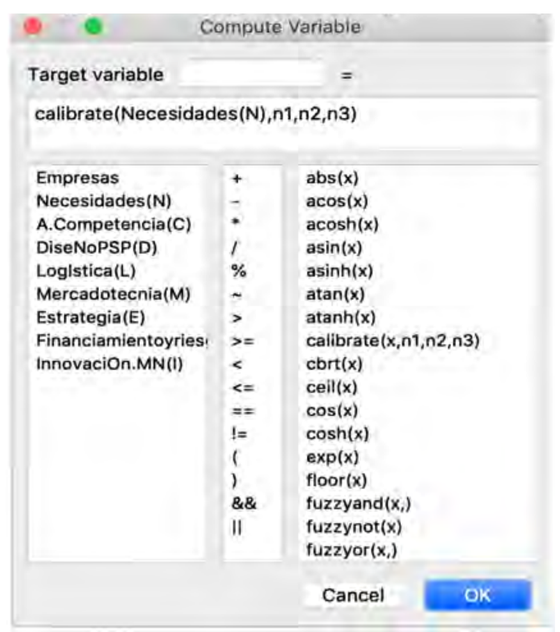
- *Adopción de innovación por el modelo de negocios* **I.** Resultado de todas las etapas en conjunto de empresas innovadoras en productos de alta tecnología.

Por lo que tiene de similar al **csQCA**, este apartado con propósitos de guía de un análisis **fsQCA,** pretende centrarse en las particularidades propias del **fsQCA** y las opciones que solamente tienen aplicación para conjuntos *difusos*, los cuales, básicamente son dos, al accesar al archivo de trabajo:

- La opción **Calibrate**, localizada al seleccionar **Variable** → **Compute** → **Calibrate**, con la que se puede calibrar una variable original, siguiendo el *método directo de calibración.* En este caso, con la variable *necesidades detectadas del consumidor,* **N**.

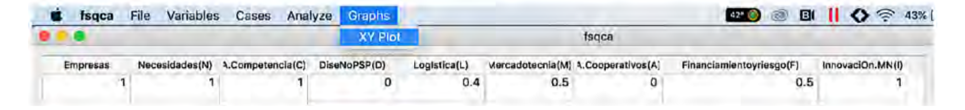



Compute Variable
Target variable
=
calibrate(x,n1,n2,n3)
Empresas
Necesidades(N)
A.Competencia(C)
DiseNoPSP(D)
Loglstica(L)
Mercadotecnia(M)
Estrategia(E)
Financiamientoyries
InnovaciOn.MN(I)
+
-
*
/
%
~
>
>=
<
<=
==
!=
(
)
&&
||
abs(x)
acos(x)
acosh(x)
asin(x)
asinh(x)
atan(x)
atanh(x)
calibrate(x,n1,n2,n3)
cbrt(x)
ceil(x)
cos(x)
cosh(x)
exp(x)
floor(x)
fuzzyand(x,)
fuzzynot(x)
fuzzyor(x,)
Cancel
OK


Compute Variable
Target variable
=
calibrate(Necesidades(N),n1,n2,n3)
Empresas
Necesidades(N)
A.Competencia(C)
DiseNoPSP(D)
Loglstica(L)
Mercadotecnia(M)
Estrategia(E)
Financiamientoyries
InnovaciOn.MN(I)
+
-
*
/
%
~
>
>=
<
<=
==
!=
(
)
&&
||
abs(x)
acos(x)
acosh(x)
asin(x)
asinh(x)
atan(x)
atanh(x)
calibrate(x,n1,n2,n3)
cbrt(x)
ceil(x)
cos(x)
cosh(x)
exp(x)
floor(x)
fuzzyand(x,)
fuzzynot(x)
fuzzyor(x,)
Cancel
OK

Además de:

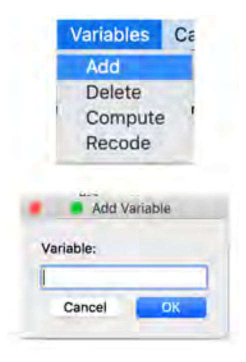
- La producción de **XY Plots,** que se encuentra en la pestaña **Graphs**.




## Matriz de datos y su preparación

El software **fsQCA 3.0**:

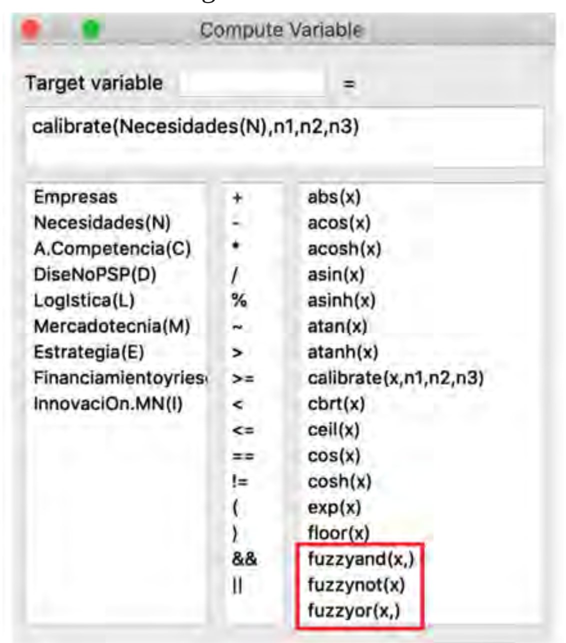
1. Permite abrir matrices de datos de varios formatos, siendo los formatos **.csv** (**EXCEL**) y **.dat** (**SPSS**) los recomendados. Igual para su guardado. Es habitual que la primera fila contenga etiquetas con los nombres de las condiciones, y también que en la primera columna se identifiquen los nombres de los casos.
2. Una vez abierta la *matriz de datos*, se pueden incorporar nuevas condiciones con **Variables** → **Add.** La opción borrar variables se encuentra en **Variables** → **Delete**.




3. Para llevar a cabo un análisis **fsQCA**, es necesario que *la matriz de datos* para análisis esté formada por valores *difusos* y, por lo tanto, *calibrada*. Si se opta por la *calibración cualitativa*, se acepta que a la *fuzzyficación* o *borrosidad* original propuesta por el investigador, no se le aplique

mayor transformación. Quizá el investigador quiera crear nuevas condiciones a partir de las ya existentes en su matriz de datos original.

4. Como se muestra en la imagen:




Se tienen **Variable → Compute**, con las funciones de opción: **fuzzyand**, **fuzzyor** y **fuzzynot** para ello. Otras operaciones aritméticas están habilitadas dentro de la casilla **Functions.**

5. Si la matriz de datos presenta *variables aún no calibradas*, una de estas funciones sirve para llevar a cabo el *método de calibración directa*.
6. Seleccionando de la lista la función **calibrate(x,n1,n2,n3)** se puede *calibrar* una condición *cuyos valores en la matriz de datos sean numéricos o de intervalos*.
7. Lo primero será indicar cuál será el resultado del proceso de calibración. Es necesario indicar un nombre para la nueva variable difusa resultante en **Target Variable**.

8. Después se selecciona, la función **calibrate** de entre todas las funciones de la lista, que aparecerá dentro del cuadro **Expression**. Esta función se compone de varios elementos: **calibrate** indica la función, **X** se refiere a la variable original que se pretende calibrar y las sucesivas comas **( , )** consideran la localización de los anclajes o valores de referencia teóricos.
9. El orden que siguen es de *total pertenencia a total exclusión* en el conjunto, y tiene que fijarse con precisión dónde se sitúan en la variable original. Es importante conocer la dirección de la calibración, es decir, si se quiere un punto *final positivo* o un punto *final negativo*, dependiendo de si la variable original mide de menor a mayor o viceversa.
10. En otras palabras, atendiendo a que el *método de calibración directa* gravita en torno a la localización *de tres valores de referencia teóricos*, y que los establece en **0.95, 0.5 y 0.05**, en ningún caso se deben escribir en esa función el número de valores *difusos* que el investigador espera conseguir (por ejemplo, **0.8, 0.6, 0.4, 0.2**).
11. Esta no es la regla; la propuesta del investigador debe cumplir el requisito de indicar en qué valores de la variable original se establece la *total pertenencia, el punto de corte y la total exclusión*.
12. Por ejemplo, si contamos con una condición que recibe el nombre *necesidades detectadas de los consumidores* de **(N)** por **23** *empresas de innovadoras de productos de alta tecnología*, que es definida como el *mínimo*, y **150**, que es el **máximo**, y contamos con argumentos teóricos para establecer los *valores de referencia teóricos* del total pertenencia en **130** casos, el *valor de referencia teórico* del total *exclusión* en **40** casos y el punto de corte en **70,** la forma de indicar estas decisiones en la casilla **Expression** es la siguiente:

**Calibrate ( X,130,70,40 )**

13. Tras esta operación, aparecerá una nueva condición en la *matriz de datos* con el nombre escrito en **Target Variable**, conservando todas las anteriores. Recuerde, que *no es necesario calibrar por el método directo todas las condiciones si hay justificación para ello*. La *matriz de*

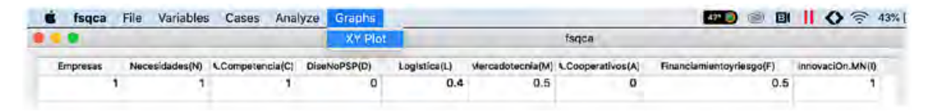
*datos* puede constar de *condiciones dicotómicas y difusas* con modelos y métodos diferentes. No obstante y es importante recordarlo, la condición que actúa como resultado tiene que ser *difusa*. De no ser *difusa,* se obtendrían dos filas de puntos horizontales paralelas en los **XY Plots** debido a que en la condición **Y,** los casos puntuarían **(0)** o **(1).** Esto dejaría desérticas las áreas interiores, repercutiendo en los parámetros de *consistencia* y *cobertura.*

**14.** De otra parte, si el método de calibración seleccionado es el *indirecto*, se sugiere primero se calcule con ayuda de **STATA** y una vez que se tengan los nuevos valores calibrados, estos se incluyan en la *matriz de datos.*

## Gráficos XY Plots

Ya se mencionó brevemente, en el segundo capítulo que los **XY Plots** son las herramientas preferentes para la inspección de los *parámetros de consistencia* y *cobertura* en conjuntos *difusos*. Con más detalle, anteriormente en este capítulo se explicó su utilidad y las pautas para su correcta lectura. Aquí se indica cómo producirlos a partir de condiciones con *valores calibrados de* **0 a 1** *(no se permiten valores continuos o multicotómicos)* al ser este un espacio dedicado a la aplicación práctica. Con el software **fsQCA 3.0**:

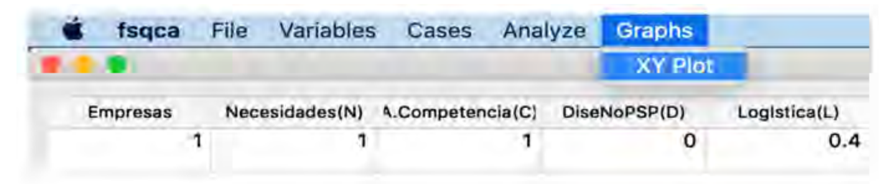
1. En la pestaña **Graphs**, seleccionar → **XYPlot**

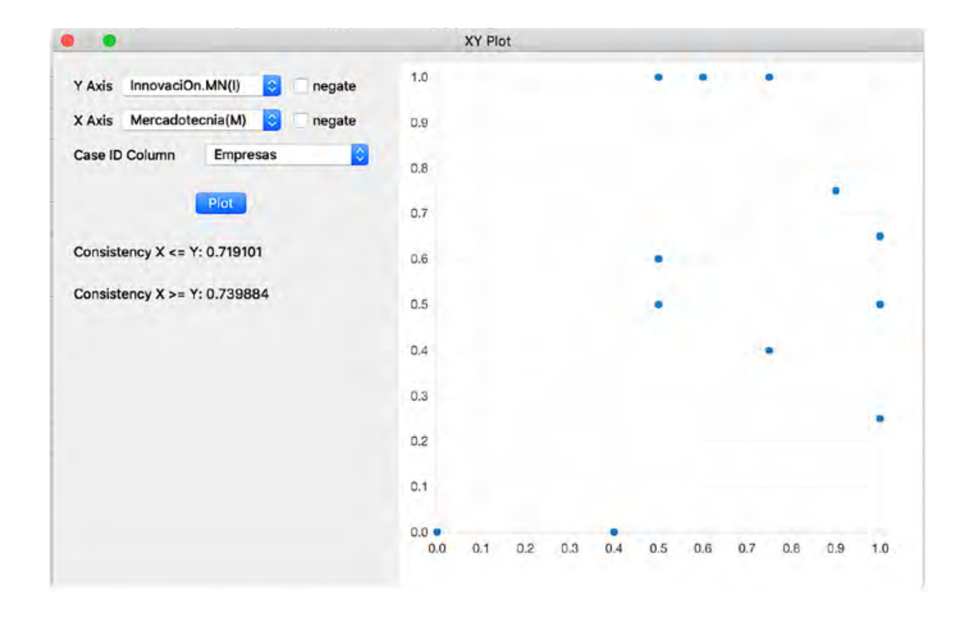



2. Lo que abre una ventana en la que aparece en el área central/derecha un **XY Plot**, seleccionar el resultado **Y Axis**, la condición causal **X Axis** y el descriptor de los casos **Empresas**. Suponga que se desea relacionar como resultado, la *adopción de la innovación por el modelo de negocios* **I**, respecto de la *mercadotecnia,* **M.**




3. Cabe respetar el lugar donde se coloca el resultado y *la condición causal* para que se pueda hacer una correcta lectura de los criterios de *necesidad y suficiencia*.
4. De nuevo, si todos los casos de **X** se encuentran *sobre o por debajo de la diagonal*, entonces hay evidencia para argumentar que **X** es una condición *necesaria* de **Y (X←Y).** Si todos los casos de **X** se encuentran *sobre o por encima de la diagonal*, entonces **X** es un *subconjunto* de **Y** y puede ser considerada una condición *suficiente* **(X → Y).**
5. Una vez seleccionadas las condiciones, el siguiente paso es pulsar el botón **Plot** que conduce a la visualización del gráfico, como se muestra en el ejemplo.






**6.** Se reportan en el gráfico las consistencia resultantes.

**(X) <= (Y): 0.719101**

**(X) >= (Y): 0.739884**

**7.** Ahora bien:

**a.** Si **X** es un *superconjunto* de **Y**, es decir, una *condición necesaria de* **Y**, localizándose mayoritariamente en el *triángulo inferior*, la *consistencia* de esa relación **Y <= X,** mientras que la casilla *superior/ izquierda* reporta la *cobertura* de esa relación de *necesidad*.

**b.** De lo contrario, si **X** es un *subconjunto* de **Y**, la casilla *superior/izquierda* ofrece el resultado del cálculo de *consistencia* de **X** como una condición suficiente de **Y**, mientras que la casilla *inferior/derech*a indica la medida de *cobertura* de dicha relación de *suficiencia.*

## Estudio de caso

El punto de partida será el previamente presentado de la *adopción de la innovación por el modelo de negocios* **I**, acorde a la **OECD (2018)**, el cual involucra las etapas de: producción, logística, mercadotecnia y arreglos cooperativos. Dicho análisis incluye **14** empresas de innovación de productos de alta tecnología hipotéticos, con los factores y valores difusos, detallados en la **Tabla 11.1.**

Este estudio parte del supuesto de la adopción de dicho modelo de innovación por negocios, en **14** empresas productoras de alta tecnología, que buscaba modernizarlas al nivel más alto de la industria, utilizando la propuesta de la **OECD (2018)**. La adopción se asume con respuestas variadas en los factores que componen al modelo, tales como los escenarios:

**Tabla 11.1. Matriz de datos difusa del caso hipotético**

| Empresas | Necesidades (N) | A. Competencia (C) | Diseño PSP (D) | Logística (L) | Mercadotecnia (M) | Estrategia (E) | Financiamiento y riesgo (F) | Innovación MN (I) |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 0.4 | 0.5 | 0 | 0.5 | 1 |
| 2 | 1 | 0.75 | 0 | 0.6 | 0.6 | 0 | 0.9 | 1 |
| 3 | 1 | 0.75 | 0 | 0.4 | 0.75 | 0 | 1 | 1 |
| 4 | 0.75 | 0.7 | 0 | 0.5 | 0.5 | 0.1 | 0.3 | 1 |
| 5 | 0.25 | 1 | 0.75 | 0.3 | 0.5 | 0 | 0.8 | 1 |
| 6 | 0.6 | 0.75 | 0 | 0.3 | 0.9 | 0 | 0.75 | 0.75 |
| 7 | 1 | 0.35 | 0.8 | 0.7 | 1 | 1 | 0.4 | 0.65 |
| 8 | 0 | 0.65 | 1 | 0.4 | 0.5 | 0.6 | 0.5 | 0.6 |
| 9 | 0.3 | 0.6 | 0.25 | 0.4 | 0.5 | 0.25 | 0.3 | 0.5 |
| 10 | 0.6 | 0.25 | 0.9 | 0.6 | 1 | 0.5 | 0.5 | 0.5 |
| 11 | 0.7 | 0.3 | 0.75 | 0.5 | 0.75 | 0.75 | 0.5 | 0.4 |
| 12 | 0.6 | 0.25 | 0.75 | 0.9 | 1 | 1 | 0.4 | 0.25 |
| 13 | 0.4 | 0 | 0.75 | 0.8 | 0.4 | 1 | 1 | 0 |
| 14 | 0.8 | 0 | 0.8 | 1 | 0 | 0.8 | 0.6 | 0 |

Fuente: Elaboración propia.

1. Algunas empresas meditaron el uso del modelo de innovación por negocios de la **OECD (2018)**.
2. Otros empresas lo rechazaron finalmente.
3. Otras empresas anunciaron su adopción pero nunca lo pusieron en práctica.
4. Otras directamente lo adoptaron, en su versión original o con alguna modificación, incluso después de fenómenos de la desaceleración económica como la causada por COVID-19 en el 2020.
5. Otras empresas que lo habían adoptado desde 2018, optaron por abandonarlo una vez aparecida la pandemia de COVID-19 en el 2020.
6. La intención de este modelo de innovación por negocios es doble:
   a. Por un lado, ser un soporte legal para normalizar las diversas políticas de innovación de la región y, por otro lado.
   b. Tratar de unificar las muchas desviaciones de conceptos, definiciones, reglas de operación, etcétera, aparecidas en la región, fruto de varias concepciones jurídicas de la innovación mezcladas con prácticas económicas múltiples, orientadas a los estímulos de la innovación.
7. La muestra del estudio incluye **14** empresas de productos de alta tecnología, donde se supone que el modelo de innovación por negocios obtiene diferentes niveles de aceptación, por ejemplo:
   a. Gran aceptación en las empresas **1, 2, 3, 4, 5.**
   b. El modelo se aceptó, con ciertas moderaciones, en las empresas **6, 7, 8.**
   c. Empresas en los que se adoptó pero nunca llegó a implementarse son la **9 y 10**
   d. En las empresas **11 y 12** se consideró la adopción pero no llegó a implementarse.
   e. Finalmente, en las empresas **13 y 14** nunca se llegó a implementar.

## Las condiciones causales que intervienen

Con el supuesto de las respuestas al modelo de innovación por negocios, basado entre otros conceptos, en la **OECD (2018)** se tienen como primeras conclusiones que el mejor diseño de investigación debe contar con condiciones causales relacionadas con la historia legal, social, económica y política tanto de la región, así como de como de cada uno de las empresas seleccionadas durante los periodos previos de las políticas de difusión del conocimiento de la innovación, hasta sus prácticas reales acordes a las políticas de estímulos otorgados institucionalmente, para el impulso de la innovación. De este modo, las condiciones que podrían llegar a explicar la causalidad son, las mencionadas líneas arriba como:

1. *Grado de necesidades detectadas del consumidor,* **N.**
2. *Análisis de la competencia,* **C.**
3. *Nivel de diseño del producto de alta tecnología a innovar,* **D.**
4. *Nivel de recursos de logística,* **L.**
5. *Mercadotecnia,* **M.**
6. *Estrategia,* **A.**
7. *Financiamiento y riesgo,* **F.**
8. *Adopción de la innovación por el modelo de negocios,* **I.**

Estas **8** *condiciones causales*, en compañía del resultado, cuentan con pleno sentido ante un estudio centrado en el porqué algunas empresas en una misma región, con características similares adoptan, en mayor o menor medida, la *adopción de la innovación por el modelo de negocios* **I,** que propone reformar sus propios sistemas económico, legales y de operación entre otros. Se observa que la naturaleza de cada una de estas condiciones es *difusa* pues es complicado definir con nitidez cuál es el grado de adopción de cada uno de los factores Tampoco es fácil de imponer una dicotomía de presencia o ausencia del *modelo de innovación por negocios* **I,** en las empresas.

Para un mayor entendimiento de las percepciones que se analizarán, se decide conceptualizar cada una de las condiciones en torno a **5** *descriptores lingüísticos* que indican:

- **a.** Total pertenencia al conjunto **(1.0),**
- **b.** Más dentro que fuera (**0.75),**
- **c.** Ni fuera ni dentro del conjunto **(0.5),**
- **d.** Más fuera que dentro **(0.25)** y
- **e.** Completamente fuera **(0.0)**.

## Criterios de asignación de valores difusos

Recuerde, estos mismos descriptores sirven para proponer una *calibración cualitativa* de los casos y la **Tabla 11.2** recoge *la propuesta teórica* detrás de la asignación de valores *difusos,* cuestión que toda investigación debe presentar.

Se destacan, las *variaciones cualitativas* entre valores que después serán empleadas como referencias para encajar los datos derivados del análisis de documentos históricos, como es el caso de este estudio, que permite completar la *matriz de datos* con valores difusos para cada uno de los casos **(Tabla 11.1)**.

El estudio, por cierto, no presenta *análisis de condiciones necesarias* para ninguno de los dos estados del resultado:

- Para cuando una empresa adopta el *modelo de adopción de la innovación por el modelo de negocios* en alguna de sus maneras, o (todos los factores en **1.0**)
- Para cuando no llega a ser implantado (todos los factores en **0.0**).

**Tabla 11.2. Criterios de asignación de valores difusos para el caso hipotético**

| Valores difusos | (N) | (C) | (D) | (L) | (M) | (E) | (F) |
|---|---|---|---|---|---|---|---|
| **1.0** | Empresa que define la planeación y ejecución de más de 3 productos de innovación en 3 años | Empresa que define quienes son los competidores más cercanos y propone acciones que se concretan más de 3 productos de innovación en 3 años | Empresa que define acciones concretas de imponer diseños innovadores de sus productos con más de 3 producto a 3 años | Empresa que percibe que su logística de insumos/ productos debe mejorarse cada año | Empresa que contrata a expertos en innovación por mercadotecnia digital y se conecta a redes sociales con 8-10 años de antigüedad | Empresa con estrategias del producto de innovación de 1 a 3 años | Empresa con accesibilidad por lograr diversificado (gobierno-socios-banca) a sus propuestas de innovación. Riesgos son considerados pero no son motivo para detenerse |
| **0.75** | Empresa que define la planeación y ejecución de 1 a 3 productos de innovación en 3 años | Empresa que define quienes son los competidores más cercanos y propone acciones que se concretan más de 1 producto de innovación en 3 años | Empresa que define acciones concretas de imponer diseños innovadores de sus productos con más de 1 producto a 3 años | Empresa que percibe que su logística de insumos/ productos debe mejorarse cada 2 años | Empresa que con sus propios recursos practica innovación por mercadotecnia digital y se conecta en redes sociales de 7 años de antigüedad | Empresa con estrategias del producto de innovación de 3 años | Empresa con accesibilidad por lograr financiamiento más que nada con socios a sus propuestas de innovación. Riesgos son considerados pero no son motivo para detenerse |
| **0.5** | Empresa que define la planeación y ejecución de sólo 1 producto de innovación en 3 años | Empresa que define quienes son los competidores más cercanos y propone acciones que se concretan en 1 producto de innovación en 3 años | Empresa que conoce las debilidades de su producto principal y ejerce medidas de diseño para innovar en 1 producto en 3 años | Empresa que percibe que su logística de insumos/ productos debe mejorarse a 3 años | Empresa que con sus propios recursos practica innovación por mercadotecnia digital y se conecta en redes sociales de 5-7 años de antigüedad | Empresa con estrategias del producto de innovación de 3 a 5 años | Empresa con accesibilidad por lograr financiamiento bancario a sus propuestas de innovación. Riesgos son considerados pero no son motivo para detenerse |

*Continuación Tabla 11.2.*

| Valores difusos | (N) | (C) | (D) | (L) | (M) | (E) | (F) |
|---|---|---|---|---|---|---|---|
| **0.25** | Empresa con actividad de búsqueda de necesidades con alta indecisión de proponer modelos innovadores | Empresa con actividad de búsqueda de determinar a sus competidores con alta indecisión de proponer modelos innovadores | Empresa que conoce la debilidad de sus productos que no se han actualizado con alta indecisión para diseñar | Empresa que percibe que su logística de insumos/ productos debe mejorarse pero que planes cambios a 5 años | Empresa que ha iniciado en la incursión de la innovación por mercadotecnia digital aunque no en redes sociales de 3-5 años de antigüedad | Empresa con estrategias del producto de innovación de 5 años | Empresa con accesibilidad por lograr financiamiento principalmente de gobierno a sus propuestas de innovación. Riesgos son considerados pero no son motivo para detenerse |
| **0.0** | Empresa con actividad de búsqueda de las necesidades del cliente pero sin mayores consecuencias | Empresa con actividad de búsqueda de las competencia pero sin mayores consecuencias | Empresa que no conoce la debilidad de sus productos que no se han actualizado y tienen bajo nivel de respuesta para diseñar | Empresa que percibe que su logística de insumos/ productos debe mejorarse pero que planea cambios a más de 5 años | Empresa que invierte en mercadotecnia pero con baja incursión en los medios de innovación por mercadotecnia digital de 1-3 años de antigüedad | Empresa con estrategias del producto de innovación de 5 a 7 años | Empresa con dificultades por lograr financiamiento a sus propuestas de innovación. Riesgos los detienen a menudo |

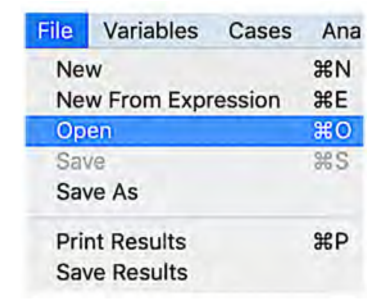
Fuente: Elaboración propia.

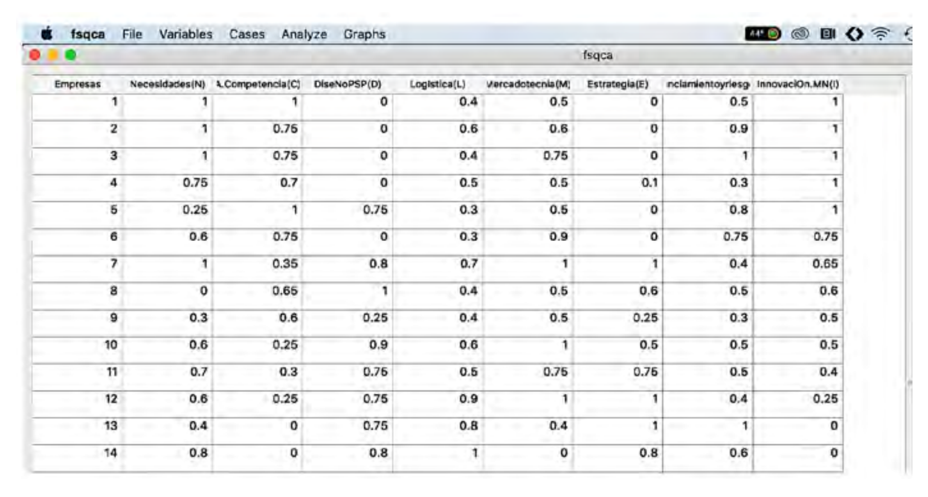
## Acceso al archivo de trabajo. Caso 1

Ya que se consideraron ambos casos muy remotos de suceder *(posibles pero no probables*). A pesar de ello, con la información proporcionada en la matriz de datos se puede realizar este análisis en **fsQCA 3.0** sin mayores obstáculos.

Así, se tiene:

1. La conversión de los datos de la **Tabla 11.1** al archivo de trabajo: **innovacionmn.cvs**, el cual, una vez abierto el software **fsQCA 3.0** se desplegará al oprimir **File → Open → innovacionmn.cvs**

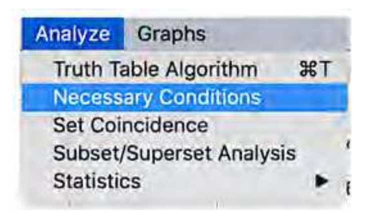


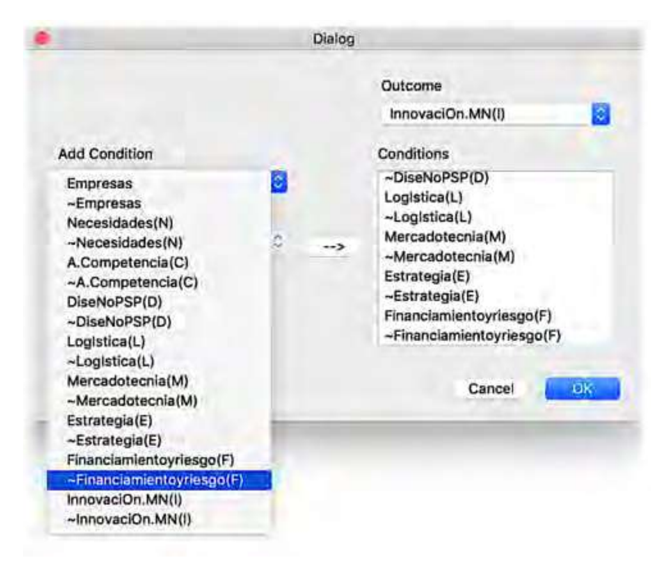

| Empresas | Necesidades(N) | A.Competencia(C) | DiseNoPSP(D) | Logistica(L) | Mercadotecnia(M) | Estrategia(E) | nclamientoyriesg | InnovaciOn.MN(I) |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 0.4 | 0.5 | 0 | 0.5 | 1 |
| 2 | 1 | 0.75 | 0 | 0.6 | 0.6 | 0 | 0.9 | 1 |
| 3 | 1 | 0.75 | 0 | 0.4 | 0.75 | 0 | 1 | 1 |
| 4 | 0.75 | 0.7 | 0 | 0.5 | 0.5 | 0.1 | 0.3 | 1 |
| 5 | 0.25 | 1 | 0.75 | 0.3 | 0.5 | 0 | 0.8 | 1 |
| 6 | 0.6 | 0.75 | 0 | 0.3 | 0.9 | 0 | 0.75 | 0.75 |
| 7 | 1 | 0.35 | 0.8 | 0.7 | 1 | 1 | 0.4 | 0.65 |
| 8 | 0 | 0.65 | 1 | 0.4 | 0.5 | 0.6 | 0.5 | 0.6 |
| 9 | 0.3 | 0.6 | 0.25 | 0.4 | 0.5 | 0.25 | 0.3 | 0.5 |
| 10 | 0.6 | 0.25 | 0.9 | 0.6 | 1 | 0.5 | 0.5 | 0.5 |
| 11 | 0.7 | 0.3 | 0.75 | 0.5 | 0.75 | 0.75 | 0.5 | 0.4 |
| 12 | 0.6 | 0.25 | 0.75 | 0.9 | 1 | 1 | 0.4 | 0.25 |
| 13 | 0.4 | 0 | 0.75 | 0.8 | 0.4 | 1 | 1 | 0 |
| 14 | 0.8 | 0 | 0.8 | 1 | 0 | 0.8 | 0.6 | 0 |

## Análisis de condiciones necesarias, con ocurrencia del fenómeno. Caso 2

2. Al oprimir **Analyse** → **Necessary conditions** se examinan todas las condiciones (y su negación) para la ocurrencia y negación de la condición *adopción de la innovación por el modelo de negocios*, **I**.




3. Para la *ocurrencia* del resultado de interés = *presencia/ mediana/ alta adopción de la innovación por el modelo de negocios* **I,** los resultados de *consistencia* y *cobertura* son los siguientes:



Analysis of Necessary Conditions

Outcome variable: InnovaciOn.MN(I)

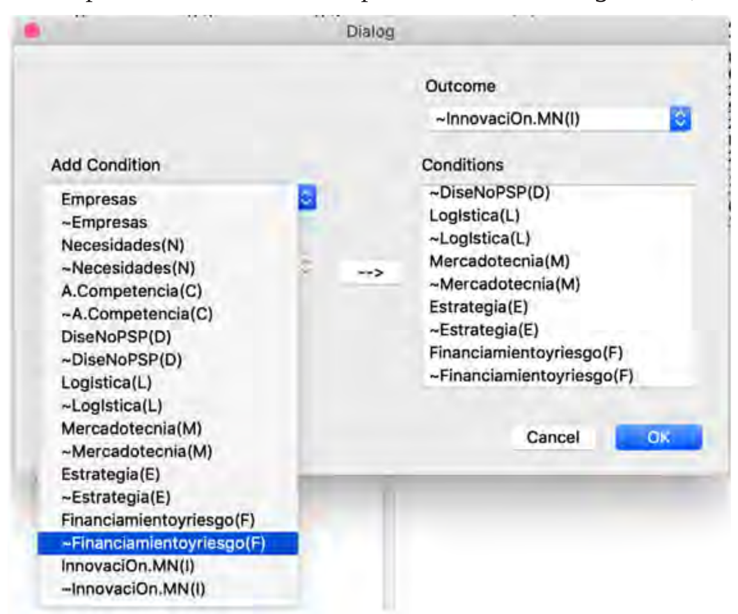
Conditions tested:

| | Consistency | Coverage |
|---|---|---|
| Necesidades(N) | 0.774567 | 0.744444 |
| ~Necesidades(N) | 0.398844 | 0.690000 |
| A.Competencia(C) | 0.832370 | 0.979592 |
| ~A.Competencia(C) | 0.416185 | 0.541353 |
| DiseNoPSP(D) | 0.393064 | 0.503704 |
| ~DiseNoPSP(D) | 0.728324 | 0.868966 |
| LogIstica(L) | 0.589595 | 0.653846 |
| ~LogIstica(L) | 0.670520 | 0.935484 |
| Mercadotecnia(M) | 0.739884 | 0.719101 |
| ~Mercadotecnia(M) | 0.404624 | 0.686275 |
| Estrategia(E) | 0.317919 | 0.458333 |
| ~Estrategia(E) | 0.843931 | 0.912500 |
| Financiamientoyriesgo(F) | 0.763006 | 0.781065 |
| ~Financiamientoyriesgo(F) | 0.520231 | 0.810811 |

## Análisis del reporte de condiciones necesarias. Caso 3

En la columna dedicada a los parámetros de *consistencia*, se observa que en ninguna de las condiciones supera el umbral de **0.95,** *que es el supuesto más conservador para el análisis de condiciones necesarias*. No obstante, una rebaja en la exigencia de este umbral es posible si, por algún motivo, se considera que no se ajusta a la naturaleza difusa del estudio. Por ejemplo, este trabajo cuenta con el supuesto de un componente altamente histórico y, además, de uso de fuentes documentadas ,empleadas para asignar valores *difusos*. La combinación de estos dos elementos puede justificar, aunque de forma *provisional*, que el valor de *consistencia* que se tome como referencia *no sea el* **0.95.** Si este fuera el caso, claro está que *no sería posible aceptar un valor extremadamente bajo*, pero posiblemente la presencia de *análisis de la competencia* **C (0.832370)** y la *ausencia/nula/baja estra-tegia* **(~ E) (0.843931)** *podrían* ser consideradas *condiciones necesarias* para el resultado de interés de la *presencia/ mediana /alta adopción de la innovación por el modelo de negocios* **I,** atendiendo también a sus valores de *cobertura*.

**4.** Para la *no ocurrencia* del resultado de interés, es decir, = *ausencia/ baja / nula adopción de la innovación por el modelo de negocios* **I,**




## Interpretación de la consistencia y la cobertura. Caso 4

Los resultados de *consistencia* y *cobertura* son los siguientes:

**Analysis of Necessary Conditions**

**Outcome variable: ~InnovaciOn.MN(I)**

**Conditions tested:**

| | Consistency | Coverage |
|---|---|---|
| Necesidades(N) | 0.710280 | 0.422222 |
| ~Necesidades(N) | 0.570093 | 0.610000 |
| A.Competencia(C) | 0.429907 | 0.312925 |
| ~A.Competencia(C) | 0.971963 | 0.781955 |
| DiseNoPSP(D) | 0.822430 | 0.651852 |
| ~DiseNoPSP(D) | 0.373832 | 0.275862 |
| LogIstica(L) | 0.925234 | 0.634615 |
| ~LogIstica(L) | 0.495327 | 0.427419 |
| Mercadotecnia(M) | 0.700935 | 0.421348 |
| ~Mercadotecnia(M) | 0.532710 | 0.558824 |
| Estrategia(E) | 0.869159 | 0.775000 |
| ~Estrategia(E) | 0.392523 | 0.262500 |
| Financiamientoyriesgo(F) | 0.803738 | 0.508876 |
| ~Financiamientoyriesgo(F) | 0.654206 | 0.630631 |

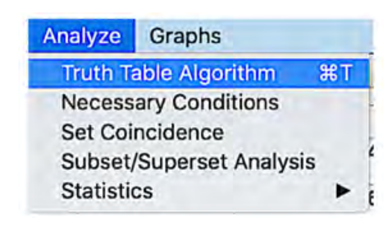
En este caso sí se observan *dos condiciones necesarias* que *superan o están cerca del umbral de consistencia*. La *ausencia/baja/nulo nivel de análisis de la competencia* **(~ C) (0.971963)** por parte de las empresas y la existencia de un *nivel de recursos logísticos* **L (0.925234)** parecen ser *dos condiciones necesarias para el* resultado de interés de la *ausencia/ baja/ nula adopción de la innovación por el modelo de negocios* **(~ I).** Esto tiene pleno sentido al considerar que la *presencia/ mediana /alta adopción de la innovación por el modelo de negocios* **I,** en cuestión, estaría controlada por la existencia de un *nivel de recursos logísticos* **L**, que se basa sólo en la *ausencia/baja/nulo nivel de análisis de la competencia* **(~ C)**.

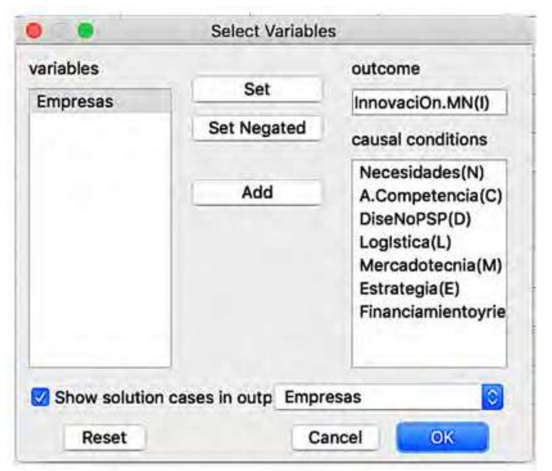
**5.** En cuanto al análisis de *condiciones suficientes* para la ocurrencia del resultado de interés, el lector debe ser consciente de una particularidad a tomar en cuenta: *el tratamiento de una configuración con una alta consistencia pero con un* **PRI** *relativamente bajo*. La buena práctica al respecto es rechazar la idea de que esa configuración pueda ser un subconjunto de **Y** pues, como dicen **Schneider y Wagemann (2012, p. 242)**, si el valor de la consistencia para **Y** y **(~ Y)** están cercanos, el valor **PRI** será bajo. No hacerlo de esta manera conduce a que la solución compute dos empresas **(7 y 12)** como ejemplos de casos en los que se adopta y rechaza la *adopción de la innovación por el modelo de negocios,* **I.** Estos dos casos son los que muestran *la particularidad de un valor de consistencia alto y un valor* **PRI** *bajo*. Si atendemos a cuál es su valor para la condición *adopción* **I,** se observa que la **empresa 7** obtiene un valor difuso de **0.65** mientras que la **empresa 12** alcanza el **0.25**.

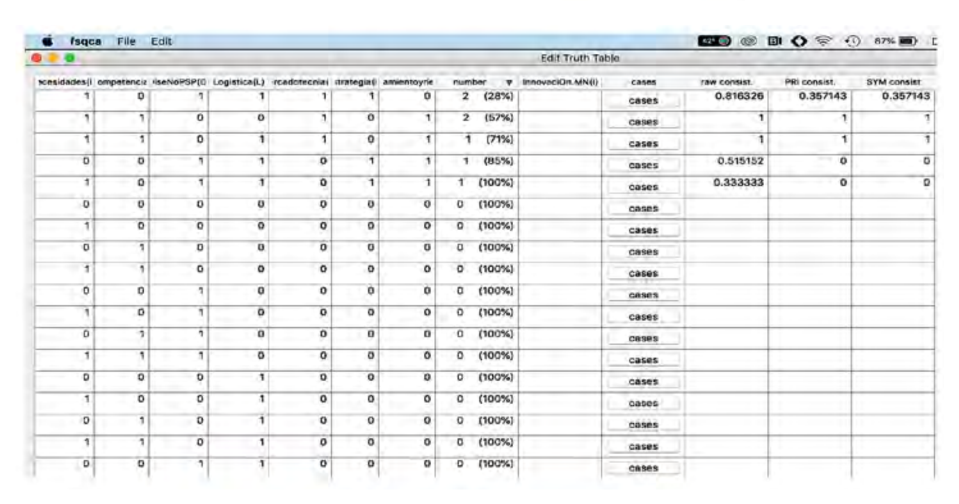
**6.** Son dos casos que bien podrían ser considerados *contradicciones,* pues en la **empresa 7** se llega *a adoptar el modelo de innovación por el modelo de negocios* **I** que en la **empresa 12** se rechaza.

## Generación de la tabla de verdad. Caso 5

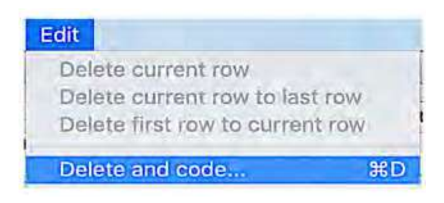
**7.** Con más detalle, se determina la *tabla de verdad*:

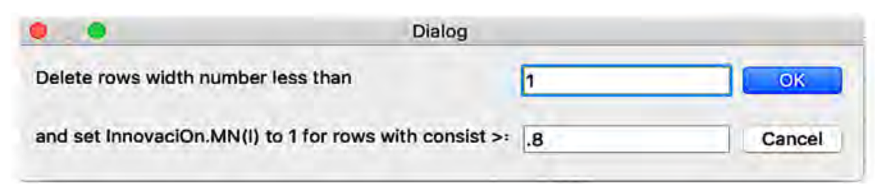


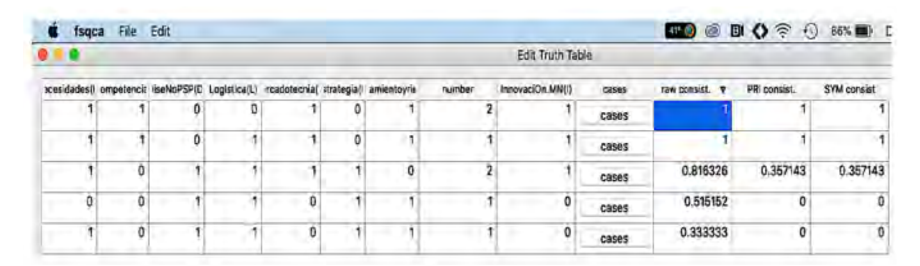




| ecesidades(I | ompetencia | iseNoPSP(D | Logistica(L) | rcadotecnia( | strategia(I | amientoyrie | number | InnovaciOn.MN(I) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 2 (28%) | | cases | 0.816326 | 0.357143 | 0.357143 |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 2 (57%) | | cases | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 (71%) | | cases | 1 | 1 | 1 |
| 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 (85%) | | cases | 0.515152 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 (100%) | | cases | 0.333333 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 (100%) | | cases | | | |
| 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 (100%) | | cases | | | |




Una vez impuestos los criterios de consistencia mayores al **0.8** (preferente **0.95**), cuestión que lleva a la ocultación de los residuales lógicos en **fsQCA 3.0,** para obtener la *tabla de verdad* a analizar:

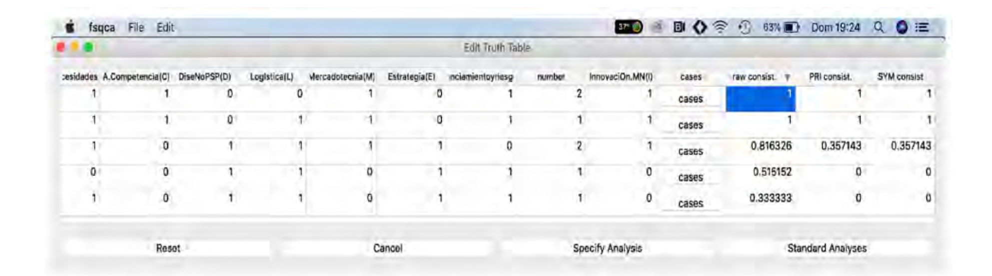



| ecesidades(I | ompetencia | iseNoPSP(D | Logistica(L) | rcadotecnia( | strategia(I | amientoyria | number | InnovaciOn.MN(I) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 2 | 1 | cases | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | cases | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | cases | 0.816326 | 0.357143 | 0.357143 |
| 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | cases | 0.515152 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | cases | 0.333333 | 0 | 0 |

## Aplicando *Standard Analyses.* Caso 6

**8.** Tras haber meditado lo anterior, quizá la decisión correcta sería interpretar que esa expresión no es susceptible de ser considerada suficiente para la *ocurrencia del resultado de interés*, indicando un **(0)** en la correspondiente casilla.

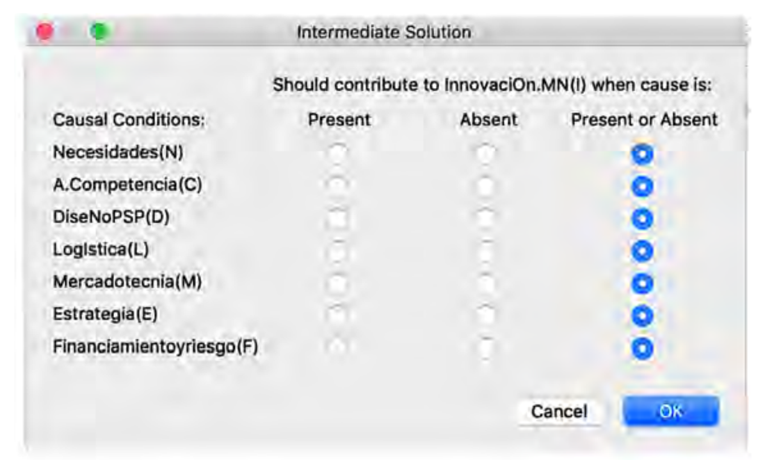
**9.** De este modo, *la primera y segunda expresión* se consideran consistentemente *subconjuntos* de **Y**, *mientras que las otras* **3** *restantes no se valoran como tal*. Así, se procede a realizar la generación de las soluciones: compleja, parsimoniosa e intermedia a partir oprimir **Standard Analyses.**



| :esidades | A.Competencia(C) | DiseNoPSP(D) | Logística(L) | Mercadotecnia(M) | Estrategia(E) | nciamientoyriesg | number | InnovaciOn.MN(I) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 2 | 1 | cases | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | cases | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | cases | 0.816326 | 0.357143 | 0.357143 |
| 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | cases | 0.515152 | 0 | 0 |
| 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | cases | 0.333333 | 0 | 0 |



Considerando, en principio, las opciones **Present / Absent** por cada una de las condiciones causales:



## Solución compleja

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/C11innovacionmn.csv
Model: InnovaciOn.MN(I) = f(Necesidades(N), A.Competencia(C), DiseNoPSP(D), LogIstica(L), Mercadotecnia(M), Estrategia(E),
Financiamientoyriesgo(F))
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.816326
                                                                                                   raw        unique
                                                                                                 coverage    coverage
consistency
                                                                                                ----------  ----------
----------
Necesidades(N)*A.Competencia(C)*-DiseNoPSP(D)*Mercadotecnia(M)*-Estrategia(E)*Financiamientoyriesgo(F)          0.421965    0.352601
1
Necesidades(N)*-A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*-Financiamientoyriesgo(F)   0.231214    0.16185
0.816326
solution coverage: 0.583815
solution consistency: 0.918182


Cases with greater than 0.5 membership in term
Necesidades(N)*A.Competencia(C)*-DiseNoPSP(D)*Mercadotecnia(M)*-Estrategia(E)*Financiamientoyriesgo(F): 3 (0.75,1),
  2 (0.6,1), 6 (0.6,0.75)
Cases with greater than 0.5 membership in term
Necesidades(N)*-A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*-Financiamientoyriesgo(F): 7 (0.6,0.65),
  12 (0.6,0.25)
```

## Solución parsimoniosa

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/C11innovacionmn.csv
Model: InnovaciOn.MN(I) = f(Necesidades(N), A.Competencia(C), DiseNoPSP(D), LogIstica(L), Mercadotecnia(M), Estrategia(E),
Financiamientoyriesgo(F))
Algorithm: Quine-McCluskey

--- PARSIMONIOUS SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.816326
                      raw       unique
                    coverage   coverage   consistency
                   ---------- ---------- ----------
Mercadotecnia(M)    0.739884   0.739884   0.719101
solution coverage: 0.739884
solution consistency: 0.719101


Cases with greater than 0.5 membership in term Mercadotecnia(M): 7 (1,0.65),
  10 (1,0.5), 12 (1,0.25), 6 (0.9,0.75),
  3 (0.75,1), 11 (0.75,0.4), 2 (0.6,1)
```

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/C1innovacionmn.csv
Model: InnovaciOn.MN(I) = f(Necesidades(N), A.Competencia(C), DiseNoPSP(D), LogIstica(L), Mercadotecnia(M), Estrategia(E),
Financiamientoyriesgo(F))
Algorithm: Quine-McCluskey

--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.816326
Assumptions:
                                                                                                  raw       unique
                                                                                                coverage   coverage
consistency
                                                                                                ---------- ----------
----------
Necesidades(N)*A.Competencia(C)*~DiseNoPSP(D)*Mercadotecnia(M)*~Estrategia(E)*Financiamientoyriesgo(F)          0.421965   0.352601
1
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F)    0.231214   0.16185
0.816326
solution coverage: 0.583815
solution consistency: 0.918182


Cases with greater than 0.5 membership in term
Necesidades(N)*A.Competencia(C)*~DiseNoPSP(D)*Mercadotecnia(M)*~Estrategia(E)*Financiamientoyriesgo(F): 3 (0.75,1),
  2 (0.6,1), 6 (0.6,0.75)
Cases with greater than 0.5 membership in term
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F): 7 (0.6,0.65),
  12 (0.6,0.25)
```

## Interpretación de las soluciones generadas. Caso 7

Recopilando la información en la **Tabla 11.3**

**Tabla 11.3. Solución *intermedia* para la ocurrencia de *adopción de la innovación por el modelo de negocios* I, sin residuales lógicos**

| | **(N * C * ~ D * M * ~E * F)** | **(N * ~ C * D * L * M * E * ~F)** |
|---|---|---|
| **Número de empresas que adoptan innovación pror el modelo de negocios explicadas** | **3** | **2** |
| **Empresas que adoptan innovación por el modelo de negocios desglosadas** | **(3, 2, 6)** | **(7, 12)** |
| Raw coverage | **0.421965** | **0.231214** |
| Unique coverage | **0.352601** | **0.16185** |
| Consistency | **1.0** | **1.0** |
| Solution coverage **0.583815** | — | — |
| Solution consistency **0.918182** | — | — |

Fuente: **fsQCA 3.0** con datos propios.

**10.** Como se ha expuesto, se tomará como base de explicación, *la solución intermedia para la ocurrencia del resultado de interés adopción de la innovación por el modelo de negocios* **I,** son las siguientes:

**Ruta 1**

**a.** *Presencia del grado de necesidades detectadas del consumidor,* **N.**
**b.** *Presencia análisis de la competencia,* **C.**
**c.** *Ausencia nivel de diseño del producto de alta tecnología a innovar,* **(~D)**
**d.** *Presencia de mercadotecnia,* **M.**
**e.** *Ausencia de estrategia,* **(~E).**
**f.** *Presencia de financiamiento y riesgo,* **F** y
**g.** *No se considera recursos de nivel de recursos de logística*, **L.**

**Ruta 2**

**h.** *Presencia de grado de necesidades detectadas del consumidor,* **N.**
**i.** *Ausencia del análisis de la competencia,* **(~C).**
**j.** *Presencia nivel de diseño del producto de alta tecnología a innovar,* **D.**
**k.** *Presencia de nivel de recursos de logística*, **L.**
**l.** *Presencia de mercadotecnia,* **M.**
**m.** *Ausencia de estrategia,* **(~ E)** y
**n.** *Presencia de financiamiento y riesgo,* **F**.

**11.** Nótese, aunque de forma indirecta, que se tienen en cuenta las condiciones que anteriormente se han considerado necesarias para la adopción de **I**. *Esto hará que la solución intermedia no contradiga estos supuestos*.

**12.** Entonces, el resultado del análisis de *condiciones suficientes* para la *ocurrencia del resultado de interés*, en su versión de *solución intermedia*, refleja que solamente hay una ruta por la que las empresas acepten la *adopción* de **I:**

**(N * C * ~ D * M * ~E * F) → (I)**

**13.** Esta solución cuenta con una *consistencia* de **1.0**, una *cobertura bruta* de **0.421965** y *cobertura única* de **0.352601**. La *consistencia general* es de **1.0** y la *cobertura de la solución* es de **0.583815** debido a que solo hay dos soluciones. Los casos que se logran explicar son los de la **empresa 3 (0.75,1), empresa 2 (0.6, 1) y empresa 6 (0.6, 0.75).**

**14.** Mientras que *la solución compleja es idéntica a la solución intermedia*, la solución *parsimoniosa* produce una posible rutas que llevan a la *adopción* de **I,** a saber:

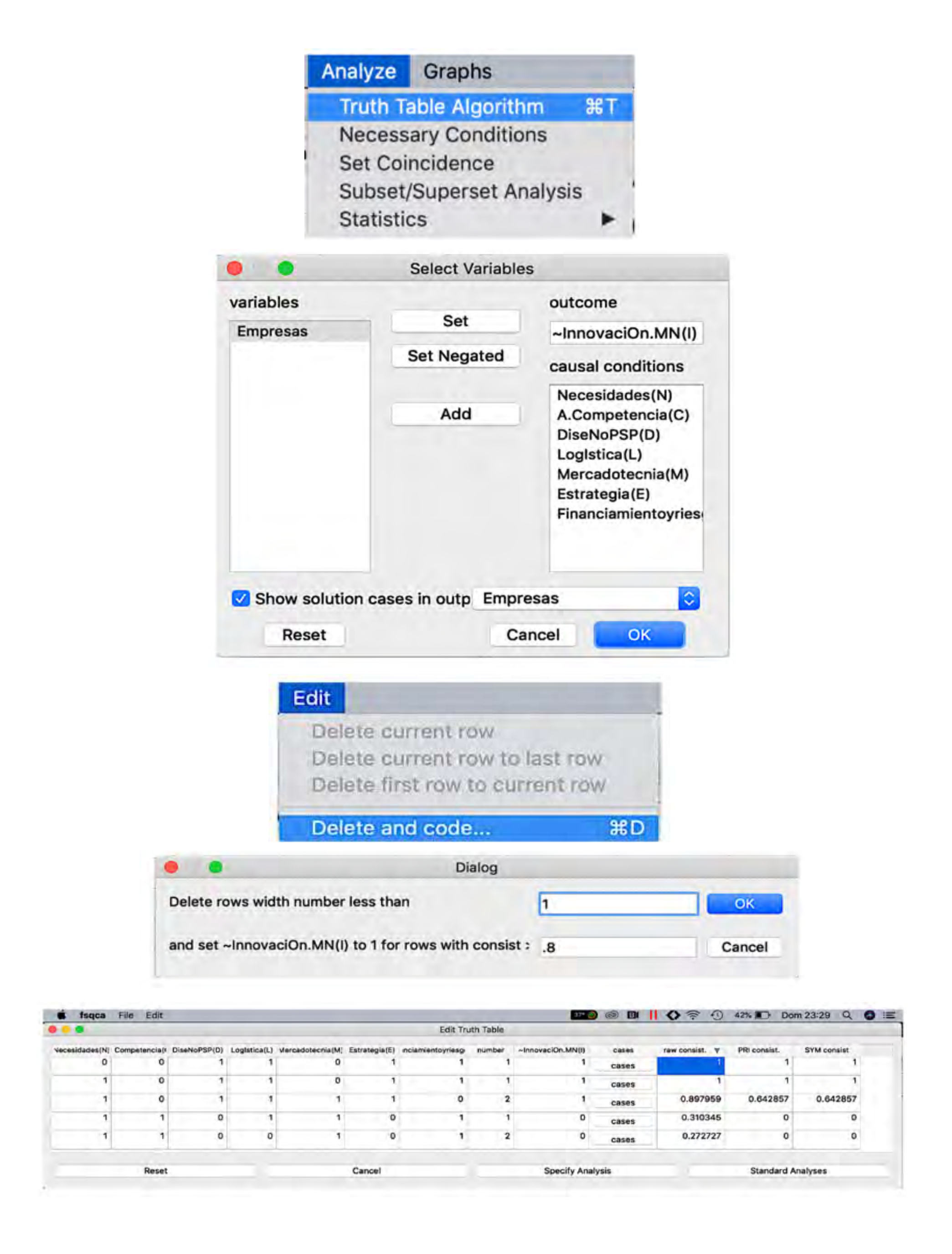
**(M) → (I)**

La *cobertura global* de esta solución es mayor que las anteriores, llegando al **0.739884**, que es prácticamente la totalidad de los casos positivos, con una consistencia general de **0.719101**. Se debe tener en cuenta que una sobredeterminación entre rutas, *genera bajos niveles de cobertura única* para cada una de ellas y la *consistencia individual* será bastante alta, por lo tanto, tiene un bajo nivel de explicación y debe evitarse.

## Generación de la tabla de verdad, con no ocurrencia del fenómeno. Caso 8

**15.** Después de explicar la ocurrencia del resultado, ¿qué explica la no *adopción de la innovación por el modelo de negocios* **I**? De nuevo, el principal asunto a tratar en el momento de indicar el criterio de *consistencia* en la *tabla de verdad* es qué hacer con los casos cuyo **PRI** *no se aproxima al valor de consistencia.* Por la experiencia anterior, el investigador conoce que estos dos casos son los de las empresas **7 y 12**, uno *no cumple el resultado y el otro lo cumple ligeramente*. Fueron tratados como *contradicciones* y así son tratados de nuevo.

**16.** La *tabla de verdad* para la *no ocurrencia del resultado.*




| Necesidades(N) | Competencia( | DiseNoPSP(D) | Loglstica(L) | Mercadotecnia(M) | Estrategia(E) | nciamientoyriesg | number | ~InnovaciOn.MN(I) | cases | raw consist. | PRI consist. | SYM consist |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | cases | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | cases | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 2 | 1 | cases | 0.897959 | 0.642857 | 0.642857 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | cases | 0.310345 | 0 | 0 |
| 1 | 1 | 0 | 0 | 1 | 0 | 1 | 2 | 0 | cases | 0.272727 | 0 | 0 |

Oprimir **Standard Analyses.**

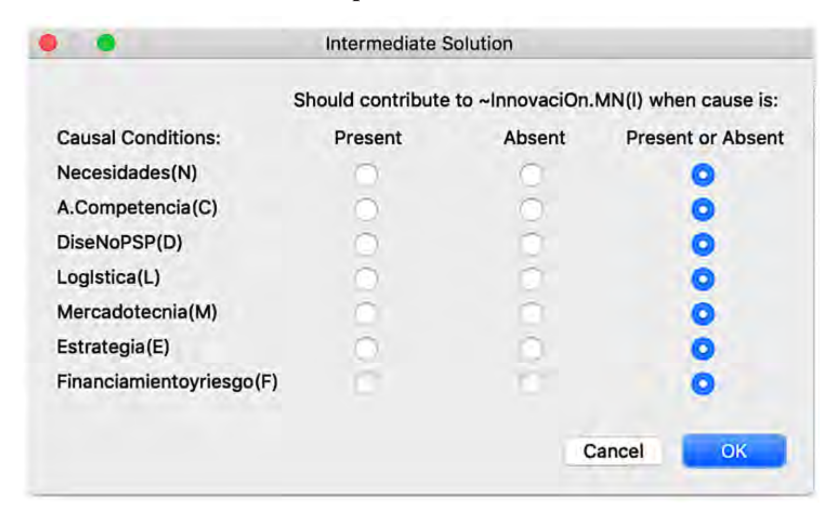
## Aplicando Standard Analyses con análisis de implicaciones principales. Caso 9

**17.** Las dos primeras configuraciones se marcan como casos positivos. Se considera que el *resto no cumplen los requisitos para ser consideradas configuraciones suficientes*. Tenga en cuenta que, cuando el algoritmo para seleccionar implicaciones principales, no puede reducir completamente la *tabla de verdad*, aparecerá la ventana **Prime Implicant Window** y el usuario debe seleccionar los implicaciones principales que se utilizarán, con base al conocimiento del marco teórico de la investigación y el conocimiento empírico sustantivo del investigador. Las implicaciones principales, son términos de producto que se producen utilizando reglas de minimización (por ejemplo, las reglas que combinan filas que difieren en una sola causa, si tienen el mismo valor de salida). Por ejemplo: **(ABC)** se combina con **(AbC)** para producir **(AC). (AC),** por lo tanto, es la implicación principal que cubre las dos expresiones Booleanas primarias **(ABC)** y **(AbC).** En otras palabras, **(ABC)** y (**AbC**) son subconjuntos de **(AC),** o **(AC)** implica **(ABC)** y **(AbC).** Sin embargo, a menudo, hay implicaciones principales más reducidos que los necesarios, para cubrir todas las expresiones primarias originales y el usuario tiene la opción de elegir entre aquellas que están vinculadas lógicamente, utilizando el gráfico de implicaciones principales **(Ragin, 1987)**. Para elegir las implicaciones principales, el programa emplea un algoritmo que intenta reducir la *tabla de verdad*, hasta que no sea posible una mayor simplificación, comenzando con los implicaciones principales esenciales (las que cubren únicamente, filas específicas en la *tabla de verdad*), que deben aparecer en la solución. Si se ejecuta el algoritmo y la tabla no se puede reducir por completo, el usuario puede seleccionar las implicaciones que se utilizarán, en función de su conocimiento teórico y empírica sustantiva**.** Aparecerá la ventana **Prime Implicant Chart**, que muestra las posibles implicaciones principales para que el usuario las elija. Cada columna en el gráfico, representa una fila diferente de la *tabla de verdad*, que está cubierta por más de una implicación principal.




La casilla del lado izquierdo describe la implicación principal que el usuario puede seleccionar. En nuestro caso, se selecciona la primera fila, como se muestra y se oprime **OK**.

**18.** Tomar esta decisión vuelve a generar un esquema complejo al que, por no tener argumentos teóricos suficientes, se marcan todas las opciones como **Present or Absent. Oprima → OK**



## Interpretando las soluciones generadas

**19.** De esta manera, son generadas las tres soluciones: *compleja, parsimoniosa e intermedia.*

### Solución Compleja

```
**********************
*TRUTH TABLE ANALYSIS*
**********************

File:  /Users/DCA/Desktop/C1innovacionmn.csv
Model: ~InnovaciOn.MN(I) = f(Necesidades(N), A.Competencia(C), DiseNoPSP(D), LogIstica(L), Mercadotecnia(M), Estrategia(E), Financiamientoyriesgo(F))
Algorithm: Quine-McCluskey

--- COMPLEX SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.897959
                                                                                              raw        unique
                                                                                            coverage    coverage    consistency
                                                                                            ----------  ----------  ----------
~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*~Mercadotecnia(M)*Estrategia(E)*Financiamientoyriesgo(F)    0.383178    0.28972     1
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F)    0.411215    0.317757    0.897959
solution coverage: 0.700935
solution consistency: 0.9375


Cases with greater than 0.5 membership in term ~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*~Mercadotecnia(M)*Estrategia(E)*Financiamientoyriesgo(F): 13
(0.6,1),
  14 (0.6,1)
Cases with greater than 0.5 membership in term
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*LogIstica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F): 7 (0.6,0.35),
  12 (0.6,0.75)
```

### Solución parsimoniosa

```
*********************
TRUTH TABLE ANALYSIS*
*********************

'ile:  /Users/DCA/Desktop/C1innovacionmn.csv
lodel: ~InnovaciOn.MN(I) = f(Necesidades(N), A.Competencia(C), DiseNoPSP(D), LogIstica(L), Mercadotecnia(M), Estrategia(E), Financiamientoyriesgo(F))
lgorithm: Quine-McCluskey

-- PARSIMONIOUS SOLUTION ---
requency cutoff: 1
onsistency cutoff: 0.897959
                      raw        unique
                    coverage    coverage    consistency
                    ----------  ----------  ----------
~A.Competencia(C)   0.971963    0.971963    0.781955
olution coverage: 0.971963
olution consistency: 0.781955


ases with greater than 0.5 membership in term ~A.Competencia(C): 13 (1,1),
  14 (1,1), 10 (0.75,0.5), 12 (0.75,0.75),
  11 (0.7,0.6), 7 (0.65,0.35)
```

## Solución intermedia



```
--- INTERMEDIATE SOLUTION ---
frequency cutoff: 1
consistency cutoff: 0.897959
Assumptions:
                                                                                                raw       unique
                                                                                                coverage  coverage  consistency
                                                                                                --------- --------- ---------
~A.Competencia(C)*DiseNoPSP(D)*Logistica(L)*~Mercadotecnia(M)*Estrategia(E)*Financiamientoyriesgo(F)    0.383178  0.28972   1
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*Logistica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F)  0.411215  0.317757  0.897959
solution coverage: 0.700935
solution consistency: 0.9375


Cases with greater than 0.5 membership in term ~A.Competencia(C)*DiseNoPSP(D)*Logistica(L)*~Mercadotecnia(M)*Estrategia(E)*Financiamientoyriesgo(F): 13
(0.6,1),
  14 (0.6,1)
Cases with greater than 0.5 membership in term
Necesidades(N)*~A.Competencia(C)*DiseNoPSP(D)*Logistica(L)*Mercadotecnia(M)*Estrategia(E)*~Financiamientoyriesgo(F): 7 (0.6,0.35),
  12 (0.6,0.75)
```

Recopilando la información en la **Tabla 11.4.**

**Tabla 11.4. Solución *intermedia* para la ocurrencia de no *adopción de la innovación por el modelo de negocios* I sin residuales lógicos**

| | **(~C * D * L * ~M * E * F)** | **(N * ~ C * D * L * M * E * ~F)** |
|---|---|---|
| **Número de empresas que adoptan innovación por el modelo de negocios explicadas** | **2** | **2** |
| **Empresas que adoptan innovación pror el modelo de negocios desglosadas** | **(13, 14)** | **(7, 12)** |
| Raw coverage | **0.383178** | **0.411215** |
| Unique coverage | **0.28972** | **0.317757** |
| Consistency | **1.0** | **0.897959** |
| Solution coverage **0.700935** | — | — |
| Solution consistency **0.9375** | — | — |

Fuente: **fsQCA 3.0** con datos propios.

**20.** La solución corresponde a la configuración

**(~C * D * L * ~M * E * F)→ (~ I)**

Los casos explicados por esta solución, que es igual a la solución compleja, son las **empresas 13 (0.6, 1) y 14 (0.6, 1).**

**21.** Para lograr explicar más casos se debe recurrir a la solución más parsimoniosa cuya *cobertura global* es de **0.971963**, aunque la *consistencia global* que consigue es de **0.781955**. La solución parsimoniosa consta de:

**(~C) → (~ I)**

La cual no es recomendable utilizar dado su escaso poder de explicación al anular los residuales.

# Glosario

Basado en la obra de **Rihoux y Ragin (2009)**, exponemos el siguiente glosario de términos, esperando le sea de utilidad al lector:

- **Case (*Caso*).** Un caso, es un objeto para el cual, los datos (mediciones) están disponibles, es decir, existen instancias empíricas para mayor referencia (por ejemplo, pozo de petróleo, vehículo terrestre, tipo de vino, paciente, país, pueblo, biblioteca, etcétera). Los casos tratados, son (o deberían ser) más conocidos (o caracterizados), que desconocidos.
- **Causal Combination (*Combinación Causal*).** Es una combinación de las *condiciones causales* o sus complementos (una *configuración*). Si hay ***k*** *condiciones causales*, entonces uno comienza **fsQCA** con $\mathbf{2}^{k}$ combinaciones causales candidatas.
- **Causal Complexity (*Complejidad Causal*).** Mientras más *condiciones causales* sean consideradas, más combinaciones de ellas pueden ocurrir, y por lo tanto, es más compleja la situación.
- **Coherence Data (*Coherencia de Datos*).** Existe coherencia de datos **(Rihoux y Ragin, 2009, p. 15)**, cuando no hay *configuraciones* contradictorias, es decir, no hay casos que sean idénticos con respecto a las *condiciones causales*, pero con diferente en el resultado (*outcome*). Verificar la coherencia de los datos, significa detectar *configuraciones* contradictorias.
- **Condition (*Condición*).** Una condición (también conocida como, variable de condición o *condición causal*, **Rihoux y Ragin, 2009, p. 182**) es una variable explicativa, que puede afectar el resultado (no es una variable independiente variable, en el sentido estadístico).
- **Configuration (*Configuración*).** Una configuración **(Rihoux y Ragin 2009, p. 82)**, es una combinación de *condiciones relevantes* para un resultado (*outcome*) dado; es análogo, a una regla de antecedentes

múltiple y puede corresponder a una, más que una, o ningún caso empírico(s).

- **Conjunctural Causation (*Causa por Conjunción*).** La causalidad por conjunción, se refiere a que la causalidad se debe no necesariamente a una sola *condición causal*, pero en cambio, sí a un grupo de *condiciones causales* que se combinan, usando la conjunción con el operador lógico **AND**.
- **Consistency (*Consistencia*).** La *consistencia* **(Rihoux y Ragin, 2009, p. 182)**, es el grado en el cual, la evidencia empírica respalda la afirmación, de que existe una relación teórica establecida basada en el marco teórico de la investigación. Una relación de subconjunto, puede indicar una condición de ser necesaria o ser suficiente; depende sobre cuál es subconjunto: de la causa (*análisis de suficiencia*) o del resultado de interés (*outcome*) (*análisis de necesidad*).
- **Contradictory Configuration (*Configuración Contradictoria*).** La configuración contradictoria **(Rihoux y Ragin, 2009, p. 182)**, es una configuración cuyo resultado (*outcome*) es 1 para algunos casos y 0 para otros casos; por lo tanto, cubre un conjunto de casos empíricos, que, aunque comparten el mismo conjunto de datos de condiciones (por ejemplo, antecedentes), muestran diferentes valores de resultado (*outcomes*) (por ejemplo, consecuentes).
- **Counterfactual Analysis (*Análisis Contrafactual*).** El análisis contrafactual, implica evaluar el resultado (*outcome*) que exhibiría un caso contrafactual si, de hecho, existiera. Es considerar el realizar un *experimento mental* o supuesto. En el análisis contrafactual, los residuos se tratan como combinaciones de no importar (*don't care*); esto da como resultado (a menudo), nuevas combinaciones causales más simples llamadas *contrafactuales*.
- **Counterfactual Case (*Caso Contrafactual*)** (ver también: ***Logical Reminder**, **Difficult Counterfactual Case, Easy Counterfactual Case.*** Un caso contrafactual, es una combinación sustancialmente relevante de *condiciones causales* que no existen empíricamente.

- **Coverage (*Cobertura*).** La *cobertura*, es una evaluación de la forma en que los términos respectivos de las fórmulas mínimas (ver *prime implicants*) cubren los casos observados. Puede haber tres tipos de *cobertura*: (1) la *cobertura bruta* (*raw coverage*), que es la proporción de los casos de resultados (*outcome*), que están cubiertos por un término dado; (2) *cobertura única* (*unique coverage*), que es la proporción de casos de resultados (*outcome*), que están cubiertos de manera única por un término dado (ningún otro término cubre esos casos); y (3) *solución por cobertura* (*solution coverage*), que es la proporción de casos que están cubiertos por todos los términos.
- **Difficult Counterfactual Case (*Caso Contrafactual Difícil*).** Los casos de *contrafactuales difíciles*, intentan eliminar una *condición causal* contribuyente de una configuración, mostrando el resultado (*outcome*).
- **Diversity (*Diversidad*).** La diversidad se refiere a si existe o no, un caso para una combinación particular de condiciones de causalidad. En aplicaciones de ciencias sociales, es muy común que no existan casos para muchas combinaciones de *condiciones causales*, y esto se conoce como diversidad limitada.
- **Easy Counterfactual Case (*Caso Contrafactual Fácil*).** Los casos *contrafactuales fáciles*, suponen que, agregar una *condición causal* a una configuración conocida para producir un resultado (*outcome*) (por ejemplo, agregar una condición D a ABC, de modo que el resultado sea ABCD) aún produciría el resultado (*outcome*).
- **Equifinality (*Equifinalidad*).** La equifinalidad, se refiere a diferentes combinaciones causales, que conducen al mismo resultado.
- **Holistic Perspective (*Perspectiva Holística*).** La perspectiva holística (Rihoux y Ragin, 2009: 6) significa que, cada caso individual es considerado como una combinación compleja de propiedades, un todo específico que no debe perderse o ser minimizado en el curso del análisis.
- **Limited Diversity (ver *Diversidad Limitada*).**
- **Logical Reminder (*Residual Lógico*).** Un *residual* lógico (también llamado caso contrafactual o caso no observado) **(Rihoux y Ragin,**

**2009, p. 182)** es una configuración (combinación de condiciones) que carece de instancias empíricas.

- **Multiple-Conjunctural Causation (*Causa Múltiple de Conjuntura*).** La frase: *causa múltiple de conjuntura* **(Rihoux y Ragin, 2009, p. 8)**, significa que diferentes trayectorias causales (cada trayectoria causal, es relevante, de manera distinta) pueden conducir, al mismo resultado (*outcome*). El término *múltiple* se refiere al número de caminos, mientras que el término, *de conjuntura*, transmite la noción de que cada trayectoria consiste en una combinación de condiciones.
- **Necessary Condition (*Condición Necesaria*).** Una condición es necesaria para un resultado (*outcome*) **(Rihoux y Ragin, 2009, p. 183)** si siempre está presente, cuando se produce el resultado. En resumen, el resultado es un *subconjunto* de la causa (la misma causa, también puede afectar otros resultados).
- **Net Effects (*Efectos Netos*).** Según **Ragin (2008, pp. 112-114)**, para estimar el efecto neto de un determinada variable independiente, el investigador compensa el impacto de las *condiciones causales* competidoras o contribuyentes, restando, de la estimación del efecto de cada variable, cualquier variación explicada, en la variable dependiente que comparte con otras variables causales. Este es el significado central de los *efectos netos*, el cálculo de la no superposición de la contribución de cada variable independiente, a la variación explicada en la variable dependiente. El grado de superposición, es una función directa de la correlación. Generalmente, cuanto mayor la correlación de una variable independiente sea con sus competidores o contribuyentes, cuanto menor es su efecto neto.
- **Outcome (*Resultado*).** Un resultado (también denominado variable de resultado) **(Rihoux y Ragin, 2009, p. 183)** es la variable a ser explicada por las *condiciones causales*; generalmente, el resultado es el foco principal de un estudio, y es análogo al *consecuente*, en una regla.
- **Prime Implicants (*Implicaciones Principales*).** Las implicaciones principales **(Rihoux y Ragin 2009, p. 183)**, son expresiones reducidas, derivadas en el curso de minimización Booleana. Cada implicación

principal, suele ser un conjunto de condiciones unidas por el operador Booleano **AND.** Un subconjunto de las implicaciones principales derivadas, constituye una fórmula mínima, la punto final de una minimización Booleana. Cada implicación principal, en una fórmula mínima, cubre una colección de *configuraciones*, de la *tabla de verdad* para un resultado dado.

- **QCA. Qualitative Comparative Analysis (*Análisis Cualitativo Comparativo*). QCA**, (Rihoux y Ragin, 2009: xix y xx) es un término general, que captura tres conjuntos principales (conjunto Booleano, conjunto multivalor y conjunto difuso) como un grupo. Eso común referirse a la versión Booleana original del **QCA** como **csQCA**, donde **cs** denota un *conjunto de datos nítido,* la versión que permite utilizar condiciones de múltiples categorías como **mvQCA**, donde **mv** denota valor múltiple (*multi-value*) y para la versión del conjunto de datos difuso de **QCA** como **fsQCA**, donde **fs** denota el conjunto de datos difuso (*fuzzy-set*).
- **Reminder (*Residual*) (Ver *Logical Reminder*).**
- **Subset Relation (*Relación de Subconjunto*).** En los conjuntos de datos nítidos, una relación de *subconjunto* **(Rihoux y Ragin, 2009, p. 184)** existe entre dos conjuntos, siempre que todos los miembros de un conjunto, estén contenidos dentro del otro conjunto. En los conjuntos de *datos difusos*, se establece que un *subconjunto* existe, en una relación entre dos conjuntos, cuando las puntuaciones de pertenencia en un conjunto, son consistentemente menores o iguales, a los puntajes de pertenencia en el otro conjunto.
- **Sufficient Condition (*Condición Suficiente*).** Una condición (o combinacines de condiciones) es *suficiente* para un resultado (*outcome*) **(Rihoux y Ragin, 2009, p. 184)** si el resultado, siempre ocurre cuando la condición está presente (sin embargo, el resultado puede ocurrir por otras razones también). En resumen, la causa es un subconjunto del resultado.
- **Truth table (*Tabla de Verdad*).** Una *tabla de verdad* **(Rihoux y Ragin, 2009, p. 184)** muestra todas las *configuraciones* (combinaciones de condiciones) unidas por el operador booleano **AND** (intersección).

# Referencias

Achen, Ch. (2005). Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. *Conflict Management and Peace Science*. 22: 327–39.

Allison, P. D. (1977). Testing for Interaction in Multiple Regression. *American Journal of Sociology* 82, pp. 144–53.

Amenta, E. y Poulsen, J. D. (1996). Social Politics in Context: The Institutional Politics Theory and Social Spending at the End of the New Deal. *Social Forces* 75, pp. 33–60.

Bandura R. (2006). A Survey of Composite Indices Measuring Country Performance: 2006 Update, United Nations Development Programme – Office of Development Studies, en: https://www.undp.org/content/dam/undp/library/corporate/Development%20Studies/measuring_country_performance_2006update.pdf

Bendix, R y Beger, B. (1959). Images of Societies and Problems of Concept Formation in Soiology, en Gross I. (Eds.). *Symposium on Sociological Theory.* NY: Harper and Row.

Bol, D. y Luppi, F. (2013). Confronting Theories Based on Necesary Relations. Making the Best of QCA Possibilities. *Political Research Quarterly 66*(1), pp. 205-210.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley Interscience.

Boswell, T. y Brown, C. (1999). The Scope of General Theory. Methods for Linking Deductive and Inductive Comparative History. *Sociological Methods and Research* 28, pp. 154–85.

Braumoeller, B. y Goertz, G. (2000). The Methodology of Necessary Conditions. *American Journal of Political Science 44*, pp. 844–58.

Byrne, D. (2002). *Interpreting Quantitative Data*. Londres: Sage.

CNET (2011). *Moore's Law to roll on for another decade*, en: https://www.cnet.com/news/moores-law-to-roll-on-for-another-decade/

Corcuf, P. (2013). *Las nuevas sociologías*. Buenos Aires: Siglo XXI.

Cronqvist, L. (2004). Presentation of TOSMANA: Adding Multi-Value Variables and Visual Aids to QCA. *Compasss Working Paper* 2004-16, en: www.compasss.org/wp.htm

Duncan, O. D. (1984). *Notes on Social Measurement*. New York: Russell Sage Foundation.

Dziallas, M. y Blind, K. (2019). Innovation indicators throughout the innovation process: An extensive literature analysis. *Technovation*. 80-81: 3-29, en: https://www.researchgate.net/publication/326256032_Innovation_indicators_throughout_the_innovation_process_An_extensive_literature_analysis/citation/download

George, A. (1979). *Case Studies and Theory Development: The Method of Structured, Focussed Comparison*. In *Diplomacy: New Approaches in History, Theory and Policy*, ed. Paul G. Lauren. New York: Free Press.

George, A. y Bennett, A. (2005). *Case Studies and Theory Development.* Cambridge, MA: MIT Press.

GII (2019). *Global Innovation Index.* INSEAD-WIPO: Soumitra Dutta, Bruno Lanvin, and Sacha Wunsch-Vincent, en: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2019.pdf

Glaser, B. y Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Weidenfeld and Nicholson.

Goertz, G. (2002). *The Substantive Importance of Necessary Condition Hypotheses*, en *Necessary Conditions: Theory, Methodology, and Applications,* ed. Gary Goertz and Harvey Starr. Nueva York: Rowman and Littlefield.

Goertz, G. (2003). *Assessing the Importance of Necessary or Sufficient Conditions in Fuzzy-sets Social Science* COMPASSS working paper WP2003-7, en: www.COMPASSS.org/wp.htm.

Goertz, G. (2006). *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press.

Goertz, G. y Starr, H. (2002). *Necessary Conditions: Theory, Methodology, and Applications.* Nueva York: Rowman and Littlefield.

Hair, J. F.; Anderson, R. E.; Tatham, R. L. y Black W. C. (1999). Análisis Multivariante. 5a. Ed. España. Prentice Hall.

HDI (2018). *Human Development Index.*United Nations Organization, en: http://hdr.undp.org/en/data

Katz, J. (1982). *Poor People's Lawyers in Transition*. New Brunswick, NJ: Rutgers University Press.

King, G.; Keohane, R.O.; Verba,S. (1994). Designing Social Inquiry. Scientific Inference in Qualitative Research, NJ: Princeton.U.Press

Kosko, B. (1993). *Fuzzy Thinking: The New Science of Fuzzy Logic*. New York: Hyperion.

Kraus, S.; Ribeiro-Soriano, D.; Schussler, M. (2017). Fuzzy-set qualitative comparative analysis (fsQCA) in entrepreneurship and innovation research –the rise of a method. *International Management Journal. 14*: 15–33 https://link.springer.com/article/10.1007/s11365-017-0461-8

Lakoff, G. (1973). Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic* 2, pp. 458–508.

Marks S. J.; Kumpel E.; Guo J.; Bartram J. y Davis J. (2018). Pathways to sustainability: A fuzzy-set qualitative comparative analysis of rural water supply programs.*Journal of Cleaner Production 205:* 789-798, en: https://www.sciencedirect.com/science/article/pii/S0959652618327410

Marx, A. y Peters, H. (2004). Win for Life: an Emprirical Exploration of the Social Consequences of Introducing a Basic Income.Compasss Working Paper Series, pp. 2004-29.

Medina, I.; Castillo, P. J.; Alamos-Concha, P. y Rihoux, B. (2017). *Cuadernos Metodológicos 56*. Madrid: Centro de Investigaciones Sociológicas.

Mejía-Trejo, J.; Sánchez-Gutiérrez, J. y Ortiz-Barrera, M. (2013). Leadership and Value Creation: The Case of the Software Developer Sector en Guadalajara, México. *Competition Forum, 11* (1), en: https://www.questia.com/library/journal/1P3-3917759181/leadership-and-value-creation-the-case-of-the-software y https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2390787

Mejía-Trejo, J.; Sánchez-Gutiérrez, J. y Maldonado-Guzmán, G. (2016). The Customer Knowledge Management and Innovation. *Contaduría y Admi-*

*nistración, 3,* (61): 456-477, en: https://www.sciencedirect.com/science/article/pii/S018610421630002X

Mejía-Trejo, J. (2019a). *Técnicas Dependientes con SPSS.* México: Alfaomega, en: https://www.alfaomega.com.mx/default/catalogo/profesional/administracion/tecnicas-dependientes-con-spss.html

Mejía-Trejo, J. (2019b). *Diseño de cuestionarios y creación de escalas.Uso del EQS en las ciencias económico-administrativas*. México: Ed. Alfaomega, en: https://www.alfaomega.com.mx/default/catalogo/profesional/administracion/dise-o-de-cuestionarios-y-creacion-de-escalas-uso-del-eqs-en-las-ciencias-economico-administrativas.html

Mejía-Trejo, J. y Bravo-Rodríguez, A. (2019). *Values-Based Innovation. Designing a model to be Applied in Management Sciences.* CUCEA-Universidad de Guadalajara-Universidad Autónoma de Barcelona, en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3512169

Mendel, J. M. y Korjani, M. (2010). Charles Ragin's Fuzzy Set Qualitative Comparative Analysis (fsQCA) Applied to Linguistic Summarization. USA: University of Southern California, en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2517981

Moore, G. (1965) Cramming More Components onto Integrated Circuits. *Electronics Magazine* Vol. 38, No. 8 (April 19, 1965).

Nunnally, J. y Bernstein, I. (1994*). Psychometric Theory*. Nueva York: McGraw Hill.

OECD (2004). *Knowledge Management Measuring. Knowledge Management in the Business Sector FIRST STEPS*. Organisation for Economic Cooperation and Development (OECD), en: https://www.oecd-ilibrary.org/science-and-technology/measuring-knowledge-management-in-the-business-sector_9789264100282-en

OECD (2008). *Hadbook on Constructing Compite Indicators. Methodology and User Guide*. Organisation for Economic Cooperation and Development (OECD), en: https://www.oecd.org/els/soc/handbookonconstructingcompositeindicatorsmethodologyanduserguide.htm

OECD (2018). Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Ed. París, Francia. Organisation for Economic Cooperation

and Developmen (OECD), en: https://www.oecd-ilibrary.org/docserver/9789264304604-en.pdf?expires=1569822203&id=id&accname=-guest&checksum=41982EA3EBE6060AEC51870D0888A774

Parasuraman, A. y Colby, Ch. L. (2015). An Updated and Streamlined Technology Readiness: TRI 2.0. *Journal of Service Research* 18 (1): 59-74.

Pawson, R. (1989). *A Measure for Measures: A Manifesto for Empirical Sociology.* New York: Routledge.

PECITI (2014-2018). Programa Especial de Ciencia, Tecnología e Innovación (2014-2018). México: CONACYT, en: http://www.siicyt.gob.mx/index.php/normatividad/nacional/631-3-programa-especial-de-ciencia-tecnologia-e-innovacion-2014-2018/file

Ragin, C. (1987). *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, C. (1994). *Constructing Social Research: The Unity and Diversity of Method*. Thousand Oaks, CA: Pine Forge.

Ragin, C. (2000). *Fuzzy-sets Social Science*. Chicago: University of Chicago Press.

Ragin, C. (2003). *Making Comparative Analysis Count.* Tucson: Mimeo, en: http://www.compasss.org/RaginDayOne.PDF

Ragin, C. (2004). From Fuzzy Sets to Crisp Truth Tables COMPASSS working paper WP2004-28, en: www.COMPASSS.org/wp.htm

Ragin, C. (2006a). The Limitations of Net Effects Thinking. In Innovative Comparative Methods for Policy Analysis: Beyond the Quantitative-Qualitative Divide, ed. Benoît Rihoux and Heike Grimm. New York: Springer.

Ragin, C. (2006b). Set Relations in Social Research: Evaluating Their Consistency and Coverage. *Political Analysis 14* (3): 291–310.

Ragin, C. (2007). *Fuzzy Sets: Calibration Versus Measurement*. Tucson: Mimeo, en: http://www.u.arizona.edu/~cragin/fsQCA/download/Calibration.pdf

Ragin, C. (2008). *Redisigning Social Inquiry. Fuzzy Sets and Beyond.* University of Chicago Press: Chicago and London.

Ragin, C. (2009a). *Reflections on casing and case-oriented research*, in D. Byrne and C. Ragin (eds.), *The SAGE Handbook of Case-Based Methods,* Chapter 31: 522-534. Thousand Oaks, CA: SAGE Publications, Inc.

Ragin, C. (2009b).*Qualitative Comparative Analysis Using Fuzzy Sets (fsQCA)* en Rihoux, B. y Ragin, C. (eds.). *Configurational Comparative Methods*. Thousand Oaks. CA: SAGE.

Ragin, C. y Pennings, P. (2005). Fuzzy Sets and Social Research.Sociological Methods an Research, 33 (4), pp. 423-430.

Ragin, C. y Rihoux, B. (2004). Qualitative Comparative Analysis (QCA): State of the Art and Prospects. *Qualitative Methods*, 2, pp. 3-13.

Ragin, C. y Sonnett. J. ( 2004). *Between Complexity and Parsimony: Limited Diversity, Counterfactual Cases and Comparative Analysis*. In Vergleichen in der Politikwissenschaft, ed. Sabine Kropp and Michael Minkenberg. Wiesbaden: VS Verlag fur Sozialwissenschaften.

Ragin, C.; Drass K. A.; Davey, S. (2007). Fuzzy-sets/Qualitative Comparative Analysis 2.0. en: http://www.socsci.uci.edu/~cragin/fsQCA/

Rihoux, B. (2006). Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research. *International Sociology 21*, pp. 679-706.

Rihoux, B. y De Meur, G. (2009). Crisp-Set Qualitative Comparative Analysis (csQCA) en Rihoux, B.; Ragin, C. (2009). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. CA: Thousand Oaks: SAGE.

Rihoux, B. y Ragin, C. (2009). *Configurational Comparative Methods. Qualitative Comparative Analysis (QCA) and Related Techniques*. CA: Thousand Oaks: SAGE.

Rogers, E. M. (1962). *Diffusion of innovations*. Nueva York: Free Press.

Rosen R. (1991). *Life Itself: A Comprehensive Inquiry into Nature, Origin, and Fabrication of Life.* EUA: Columbia University Press.

Saisana, M.; Tarantola, S. y Saltelli A. (2005). Uncertainty and sensitivity techniques as tools for the analysis and validation of composite indicators, *Journal of the Royal Statistical Society A*, *168*(2), pp. 307-323.

Saltelli A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research* 81, pp. 65-77.

Sartori, G. (1970). Concept Misformation in Comparative Politics. *American Political Science* Review 64, pp. 1033-53.

Sartori, G. (2011) *Cómo hacer ciencia política.* Madrid: Taurus.

Schneider, C. Q. y Grofman, B. (2006). *It Might Look Like a Regression… But It's* Not! *An Intuitive Approach to the Presentation of QCA and fs/QCA Results*. Mimeo. Budapest: Irvine, en: http://www.compasss.org/SchneiderGrofman2006.pdf.

Schneider, C. Q. y Rohlfing, I. (2013). Combining QCA and Process Tracing in Set-Theoretic Multi-Method Research.*Sociological Methods & Research, 42*(4): 559-597.

Schneider, C. Q. y C. Wagemann, C. (2010). Standards of Good Practice in Qualitative Comparative Analysis. *Comparative Sociology 9*: 397-418.

Schneider, C. Q. y Wagemann, C. (2012). *Set Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis.* Cambridge: Cambridge University Press.

Sharpe, A. (2004). *Literature Review of Frameworks for Macro-indicators*, Ottawa, CAN: Centre  for the Study of Living Standards.

Smithson, M. y Verkuilen, J. (2006). *Fuzzy Set Theory.* Thousand Oaks, CA: Sage.

Smithson, M. (1987). *Fuzzy Set Analysis for the Behavioral and Social Sciences.* New York: Springer-Verlag. Princeton, NJ: Princeton University Press.

Trillas, E. (1992). *La Lógica Borrosa,* en Trillas, E. Y Gutiérrez , J. (eds.). *Aplicaciones de la lógica borrosa*. Madrid: CSIC.

Trillas, E.; Alsina, C. y Terricabras, J. M. (1995). *Introducción a la lógica borrosa.* Barcelona: Ariel.

Verkuilen, J. (2005). Assigning Membership in a Fuzzy Set Analysis. *Sociological Methods & Research 33* (4): 462.

Vis, B. (2010). *Politics of Risk-Taking.Welfare State reformin Advanced Democracies*. Amsterdam: Amsterdam Univeristy Press.

Wagemann, C. y Schneider, C. Q. (2010). Qualitative Comparative Analysis (QCA) and Fuzzy-sets: Agenda for a Research Approach and a Data Analysis Technique. *Comparative Sociology* (9): 376-96.

Walker, H. y Cohen, B. (1985). Scope Statements: Imperatives for Evaluating Theory. *American Sociological Review, 50:*288-301.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control, 8*(3): 38–53.

Zadeh, L. A. (1968). Fuzzy Algorithms. *Information and Control. 12*(2): 99-102.

Zadeh, L. A. (1972). A Fuzzy-sets-Theoretic Interpretation of Linguistic Hedges. *Journal of Cybernetics* 2(3), pp. 4–34.




Zapopan, Jal. a 5 de Septiembre de 2022

**Dictamen de Obra AMIDI.DO.20220830.fsQCA**

Los miembros del equipo editorial de la Academia Mexicana de Investigación y Docencia en Innovación (**AMIDI**), ver:
https://www.amidibiblioteca.amidi.mx/index.php/AB/about/editorialTeam
se reunieron para atender la invitación a dictaminar el libro:

***ANÁLISIS CUALITATIVO COMPARATIVO TOMO II. DIFUSO (fsQCA).Teoría y Práctica***

Cuyo autor de la obra es el **Dr. Juan Mejía Trejo**

Dicho documento fue sometido al proceso de evaluación por pares doble ciego, de acuerdo a la política de la editorial, para su dictaminación de aceptación, ver: https://www.amidibiblioteca.amidi.mx/index.php/AB/procesodeevaluacionporparesenciego

Los miembros del equipo editorial se reúnen con el curador principal del repositorio digital para convocar:

1. Que el comité científico, de forma colegiada, revise los contenidos y proponga a los pares evaluadores que colaboran dentro del comité de redacción, tomando en cuenta su especialidad, pertinencia, argumentos, enfoque de los capítulos al tema central del libro, entre otros.

2. Se invita a los pares evaluadores a participar, formalizando su colaboración.

3. Se envía así, el formato de evaluación para inicio del proceso de evaluación doble ciego a los evaluadores elegidos de la mencionada obra.

4. El comité científico recibe las evaluaciones de los pares evaluadores e informa a el/la (los/las) autor(es/as), los resultados a fin de que se atiendan las observaciones, el requerimiento de reducción de similitudes, y recomendaciones de mejora a la obra.

5. La obra evaluada, consta de:

   **introducción, 8 capítulos, glosario y referencias en 392 páginas**

Av. Lázaro Cárdenas 3454 int. 6,
Col. Jardines de los Arcos, C.P. 44500,
Guadalajara, Jalisco, México
Tel. Oficina. 33 3560 7860/ Cel. 3312809887
editorial@scientiaetpraxis.amidi.mx




**6.** El desglose de su contenido, de describe a continuación:

| Capítulo | Páginas |
|---|---|
| Introducción | 1-4 |
| Capítulo 4. Análisis Cualitativo Comparativo de Datos Difusos (fsQCA) | 5-136 |
| Capítulo 5. Evaluación del conjunto de datos fsQCA | 137-172 |
| Capítulo 6. La calibración y su importancia en fsQCA | 173-230 |
| Capítulo 7. Pensamiento configracional en fsQCA | 231-254 |
| Capítulo 8. Efectos netos en fsQCA | 255-272 |
| Capítulo 9. Efectos netos vs. Configuraciones en fsQCA | 273-300 |
| Capítulo 10. Guía de análisis csQCA | 301-344 |
| Capítulo 11. Guía de análisis fsQCA | 345-378 |
| Glosario | 379-384 |
| Referencias | 385-392 |

**7.** Una vez emitidas las observaciones, el requerimiento de reducción de similitudes, y recomendaciones de mejora a la obra por los evaluadores y todas ellas resueltas por el/la (los/las) autor(es/as), el resultado resalta que el contenido del libro:

**a.** Reúne los elementos teóricos actualizados y prácticos desglosados en cada uno de sus capítulos.
**b.** Los capítulos contenidos en la obra, muestran claridad en el dominio del tema, congruencia con el título central del libro, y una estructura consistente
**c.** Se concluye finalmente, que la obra dictaminada, puede fungir como libro de texto principal o de apoyo tanto para estudiantes de licenciatura como de posgrado.

**8.** Por lo que el resultado del dictamen de aceptación de la obra fue:

**FAVORABLE PARA SU PUBLICACIÓN**

Sirva la presente para los fines que a los Interesados convengan.

Atentamente

Dr. Carlos Omar Aguilar Navarro.
ORCID: https://orcid.org/0000-0001-9881-0236
Curador Principal AMIDI.Biblioteca
AMIDI